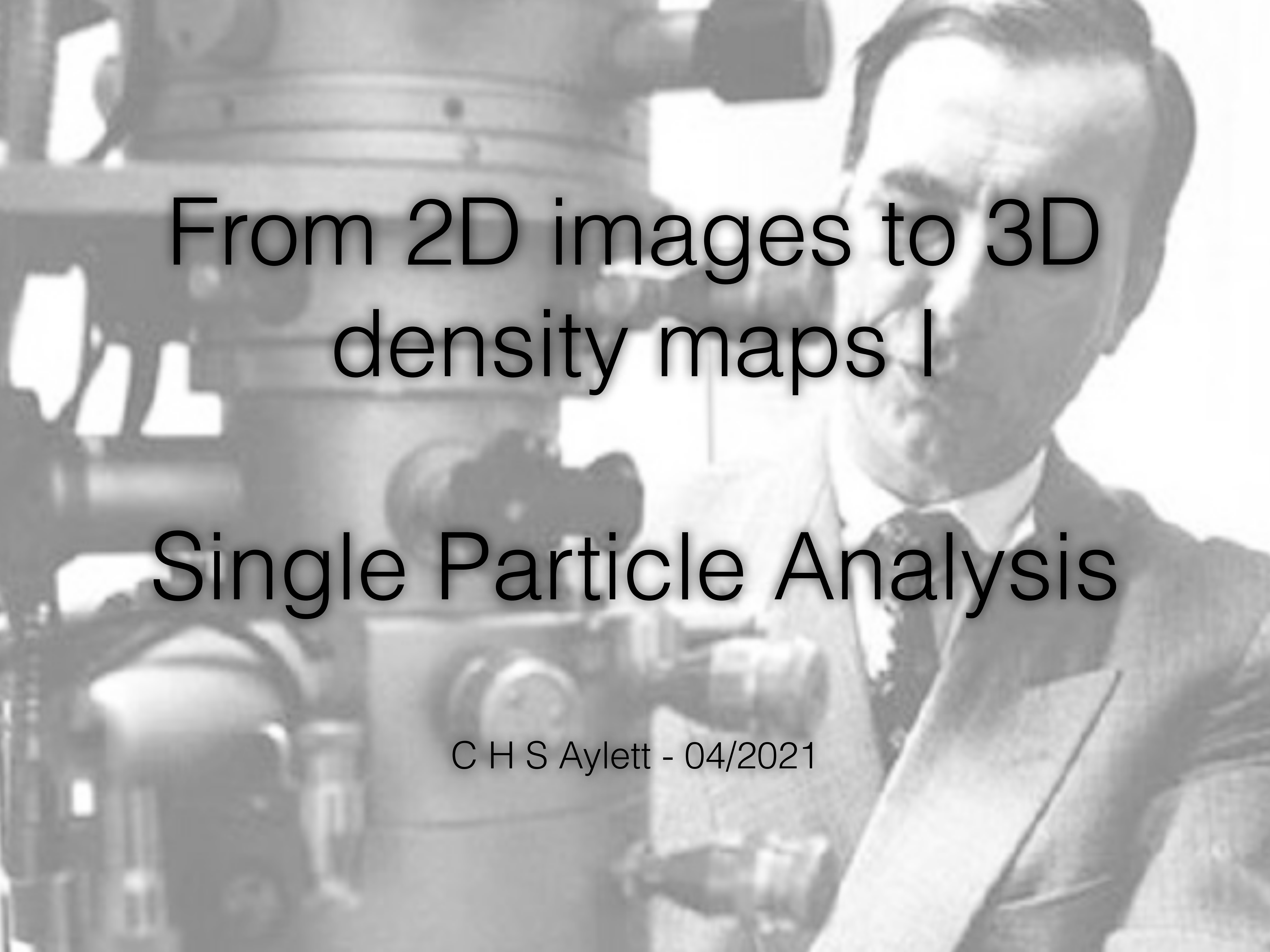
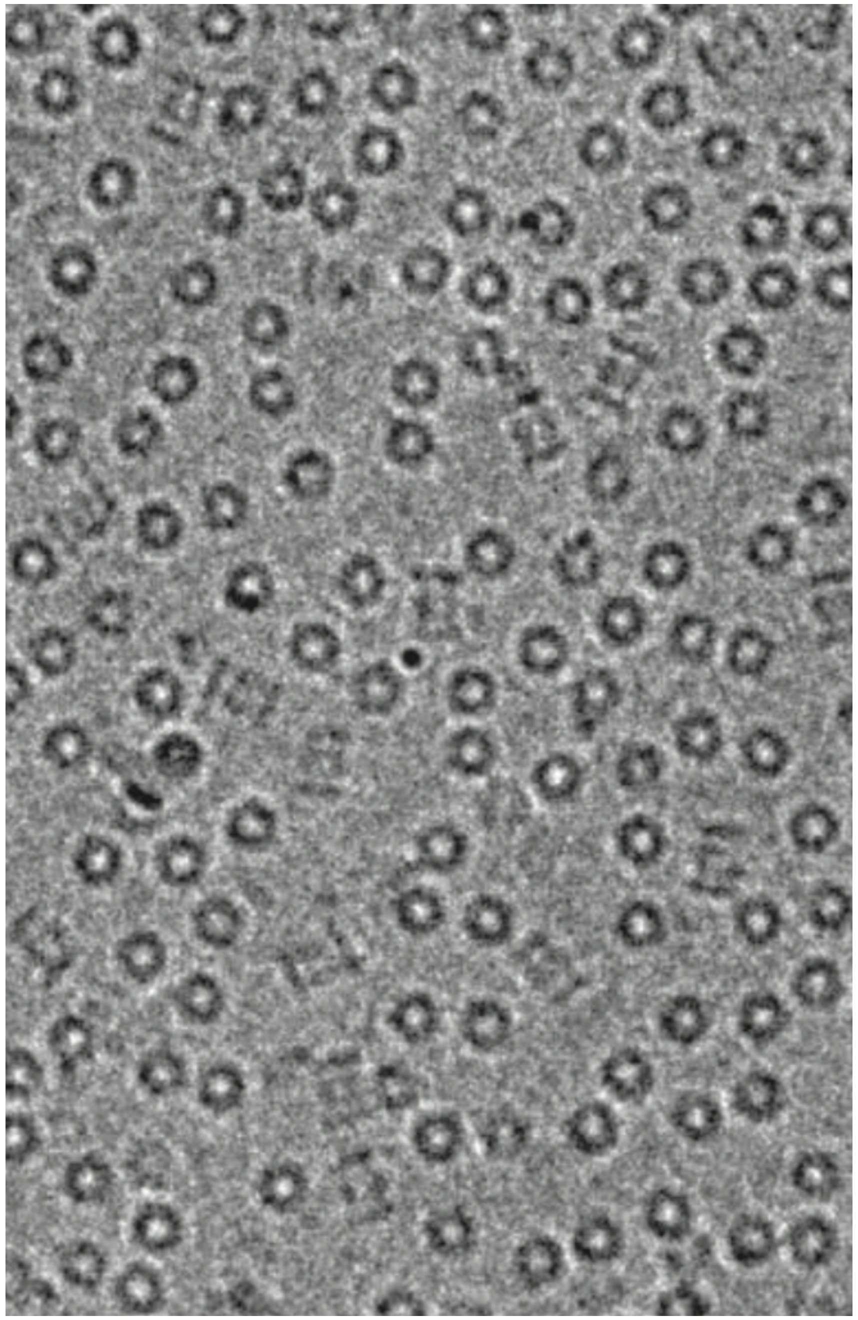


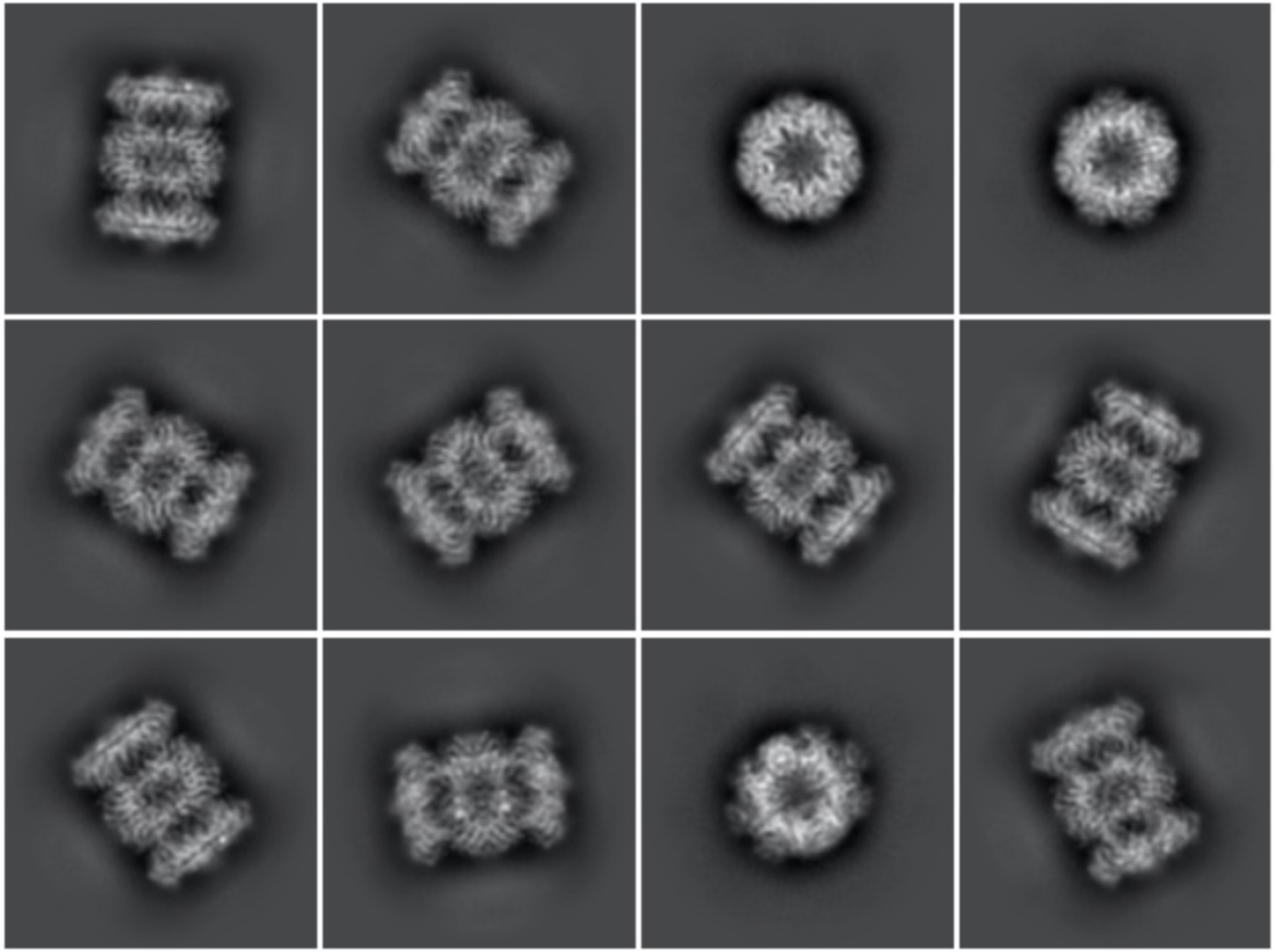
From 2D images to 3D density maps I

Single Particle Analysis

C H S Aylett - 04/2021

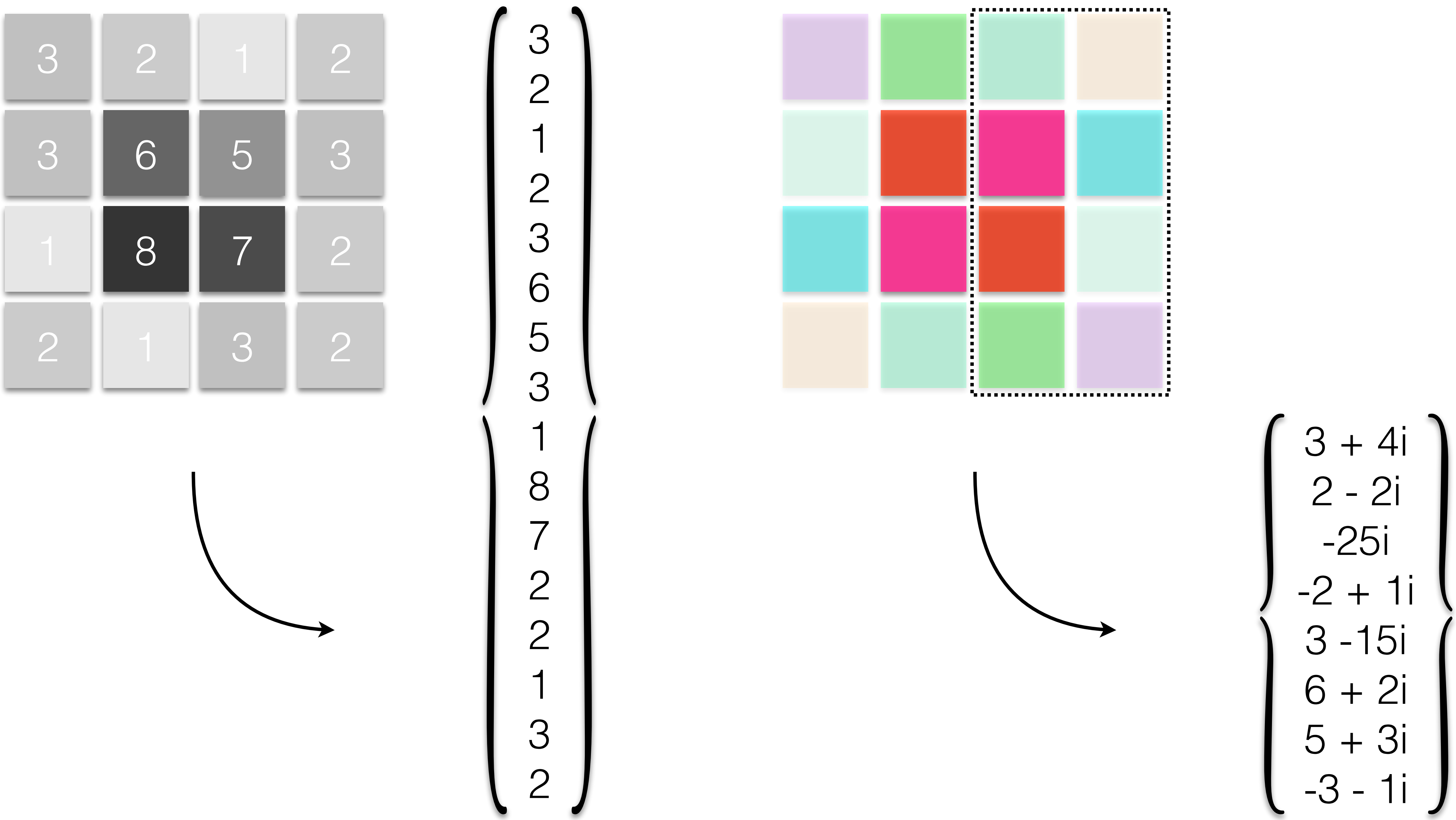
Single particle analysis - 2D





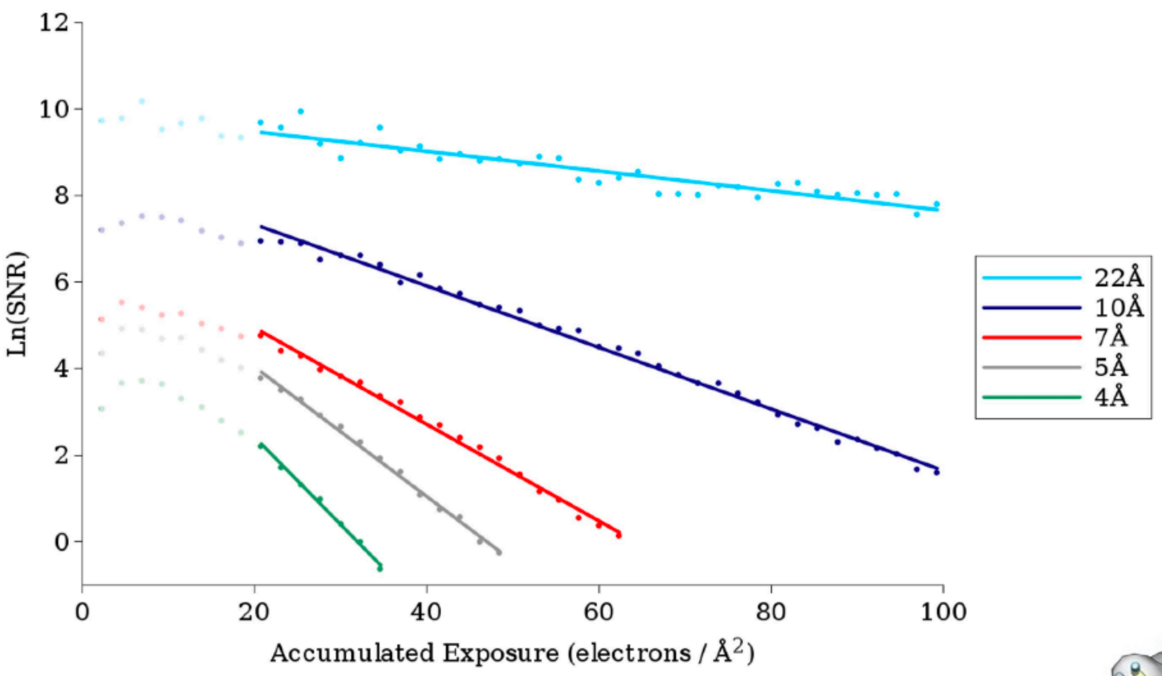
2D? We want pictures - we take pictures … right?

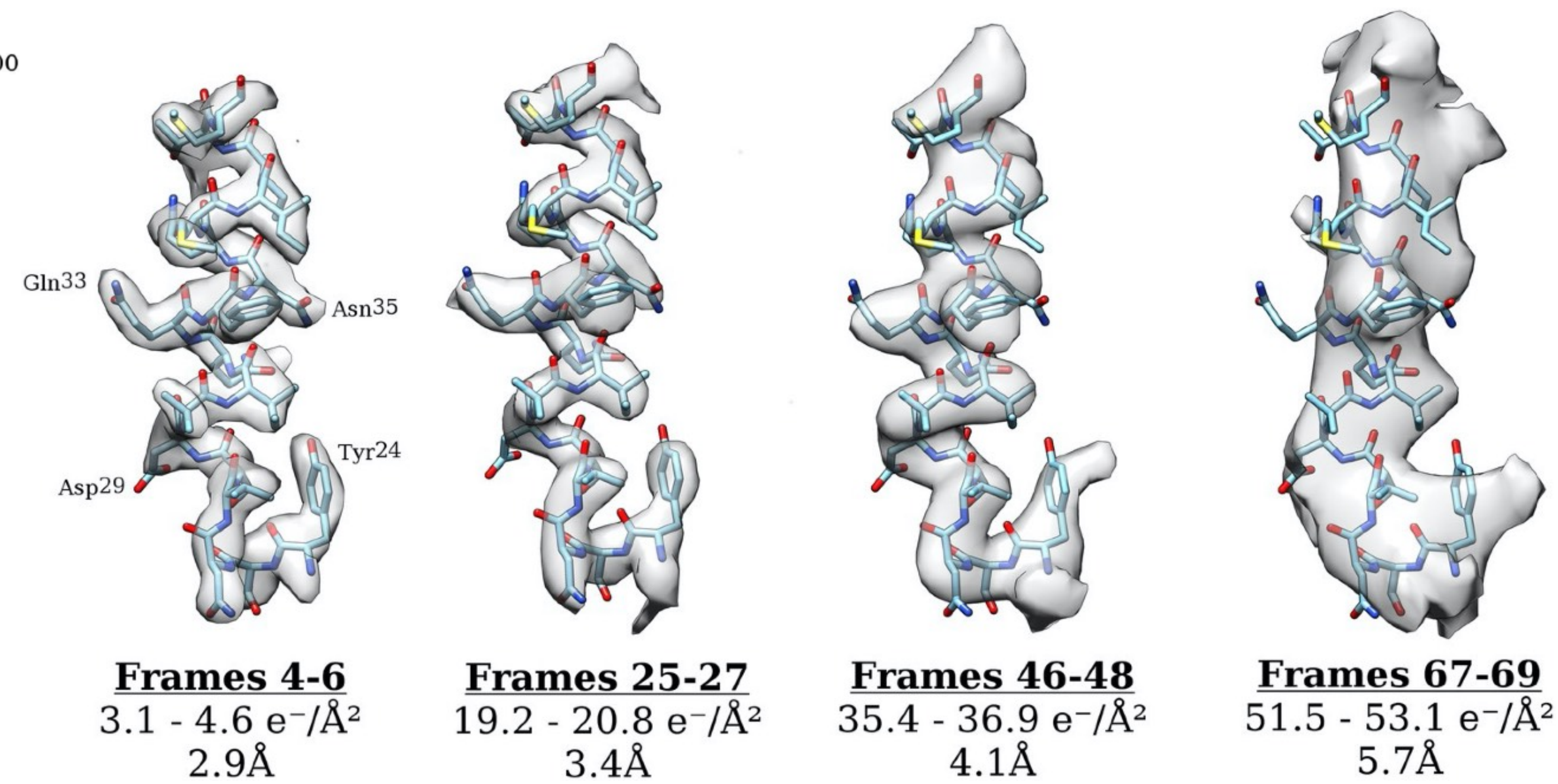
Image



X / Y grid of electron impact counts per frame

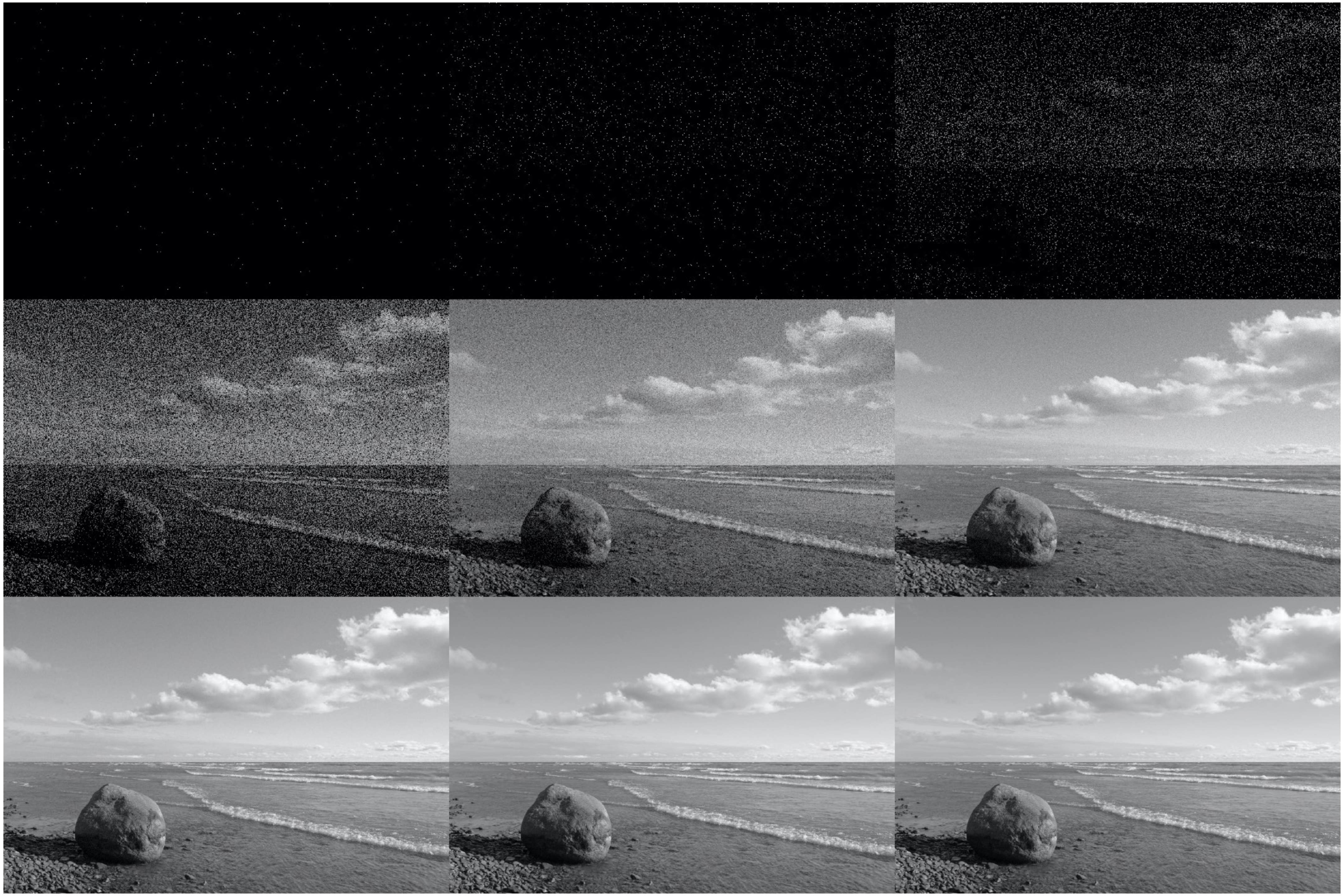
Dose





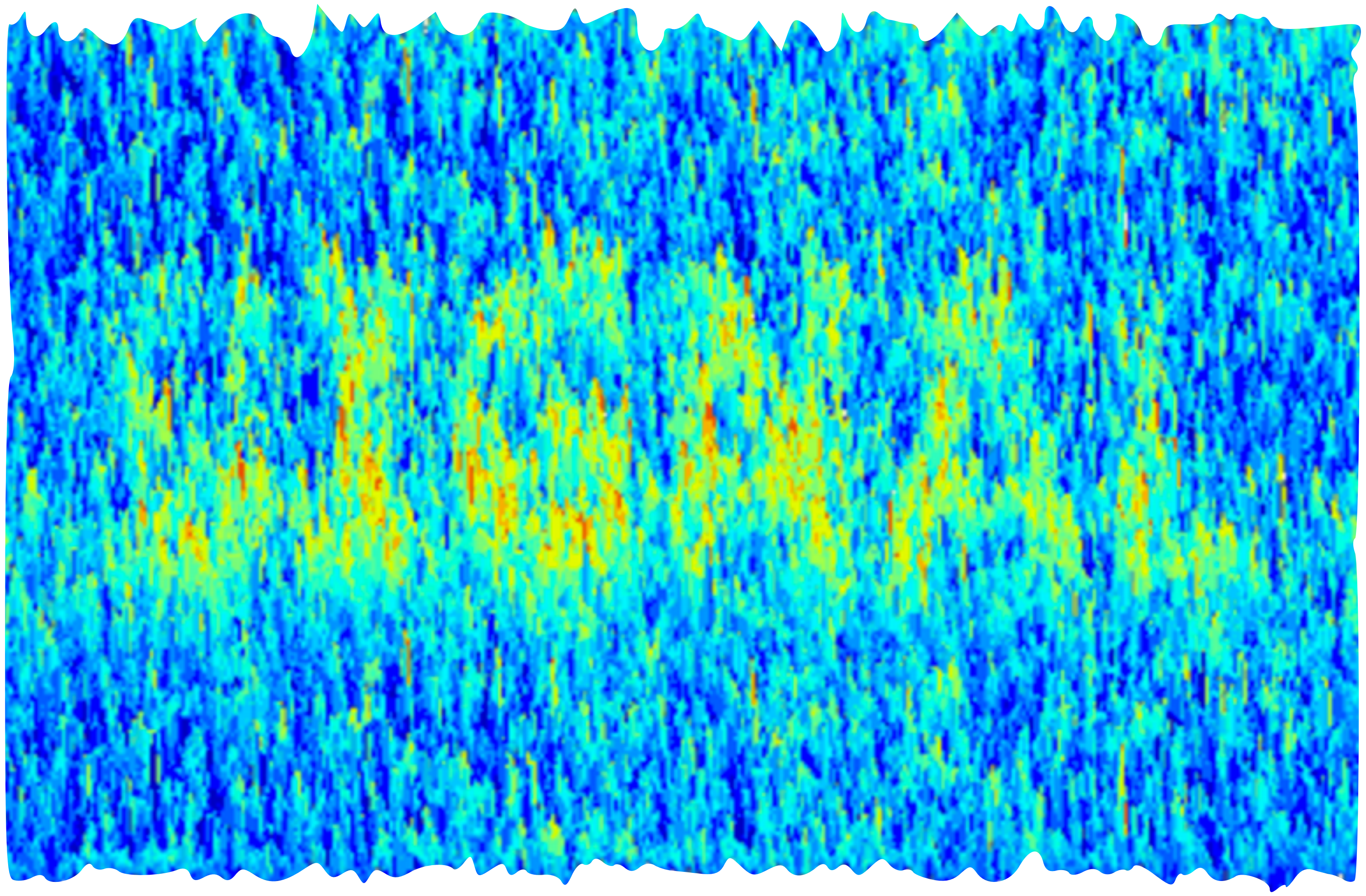
Limited to a bare minimum by radiation damage

Noise



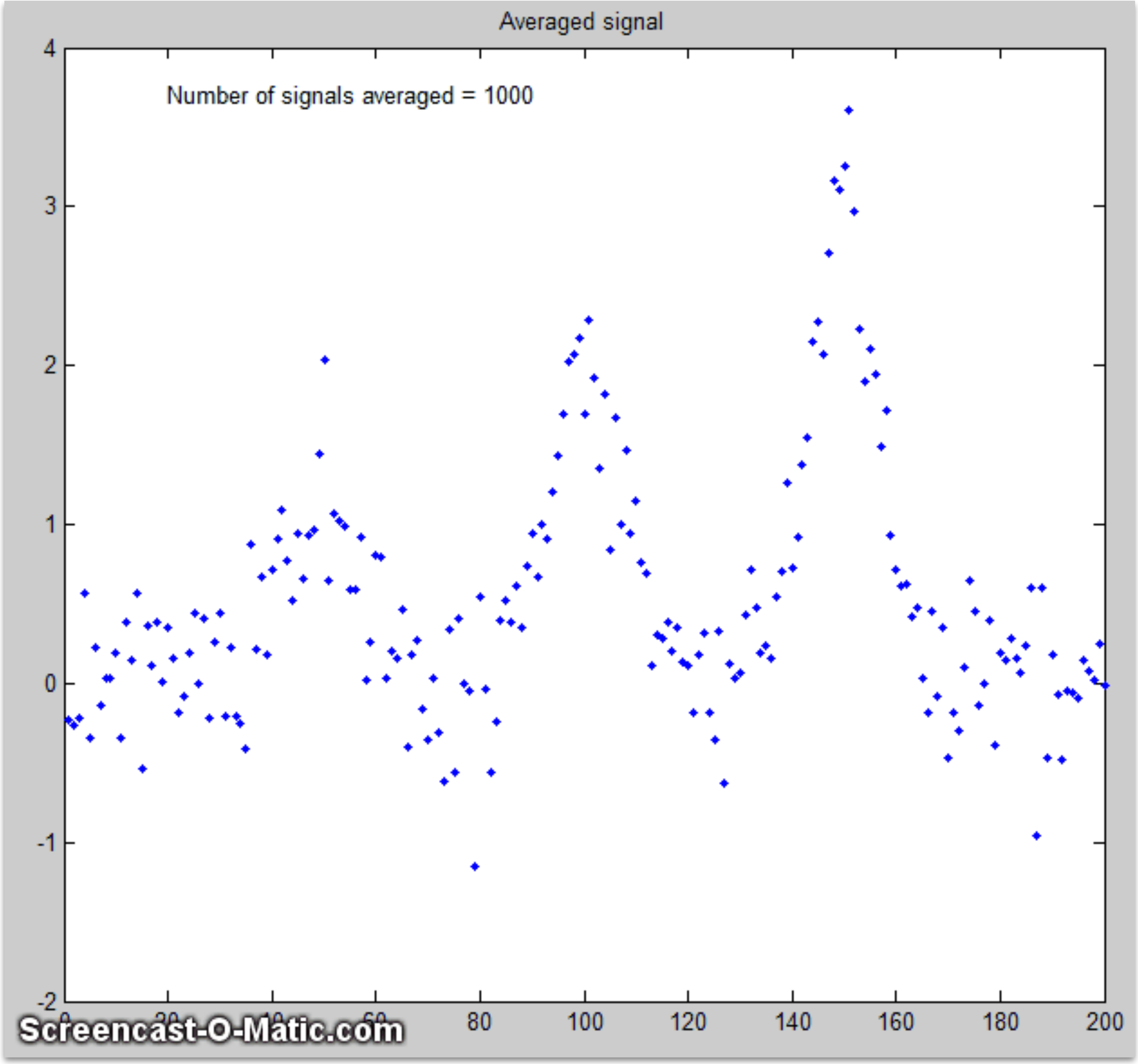
Assumed Gaussian - mostly caused by low dose

Signal



Unknown or assumed - identified by comparison

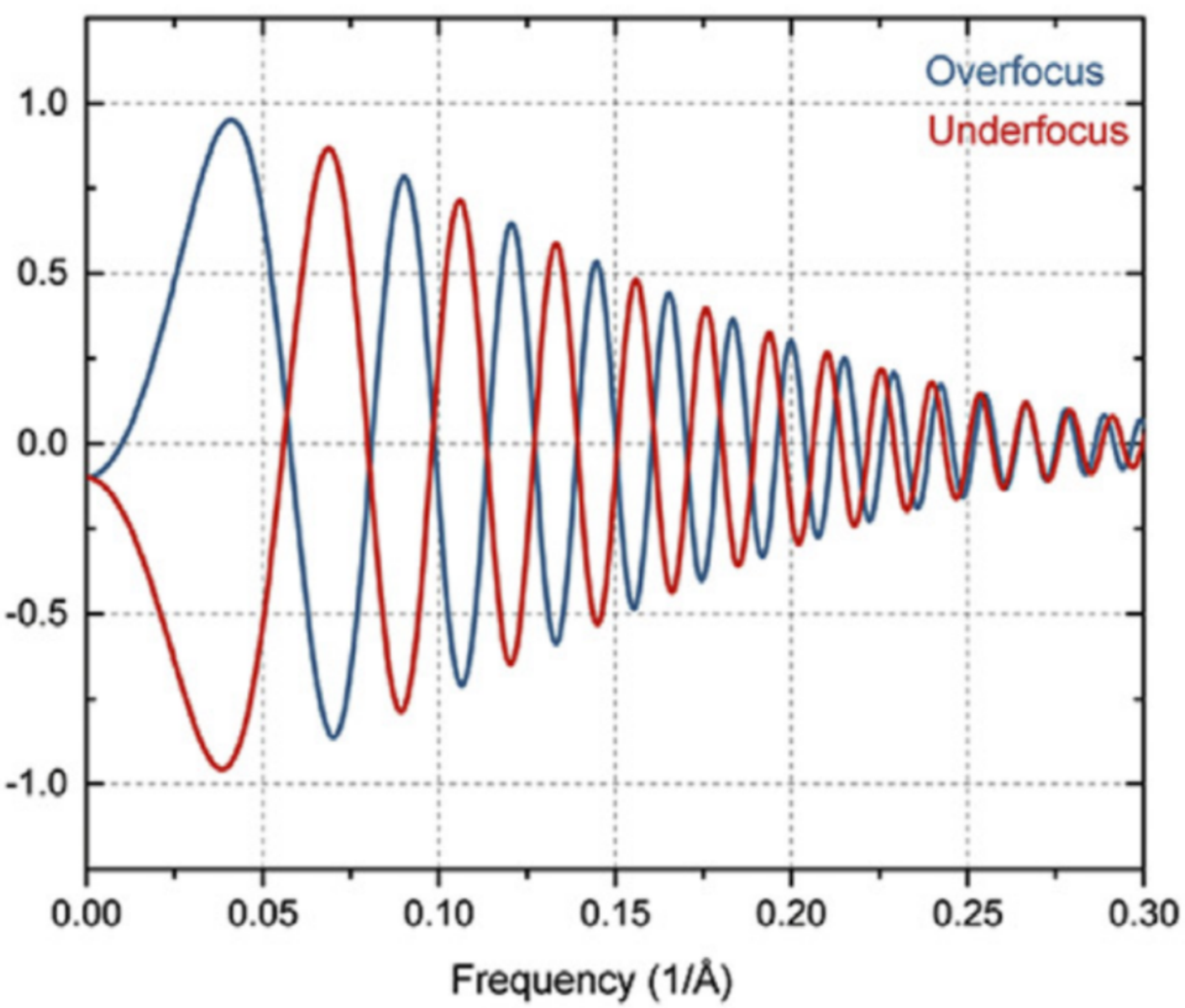
SNR

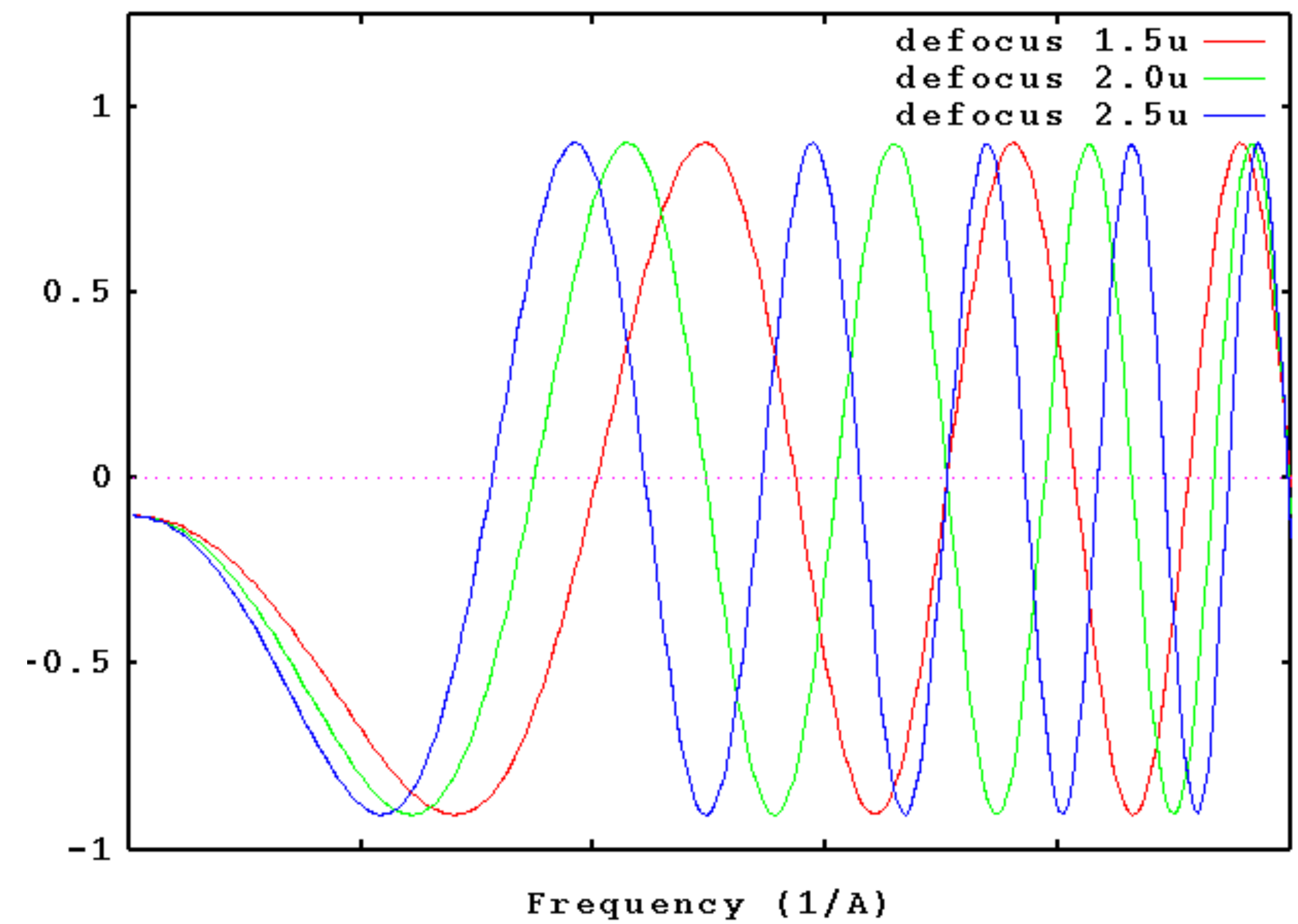


Increases with square root of number of images

CTF and PSF affect averaging

Fourier image = Signal • CTF • PSF + Noise

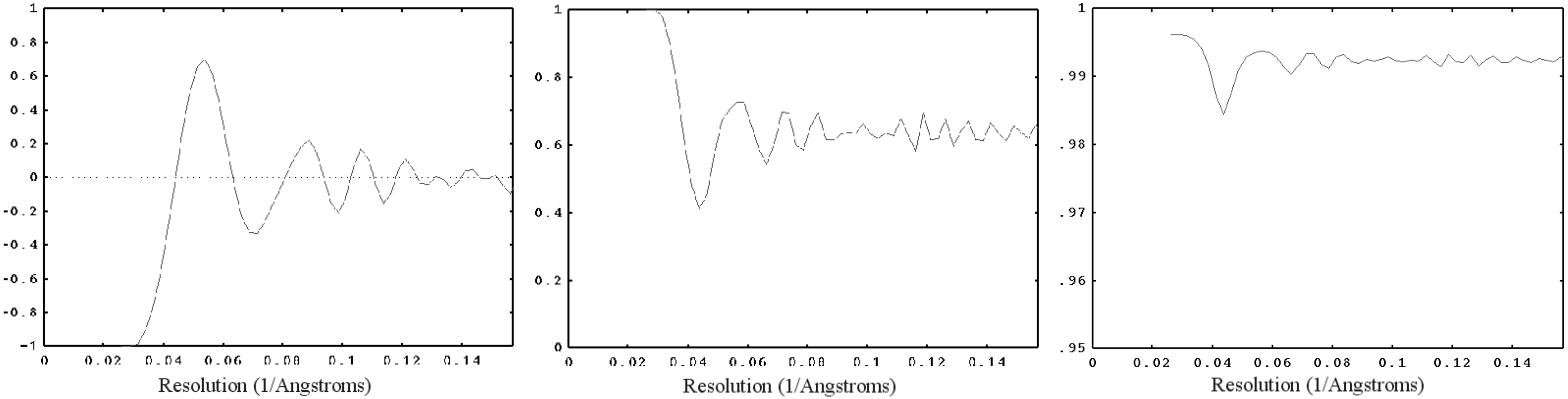




Images cannot be averaged due to CTF

The Wiener filter

Fourier space averages of …

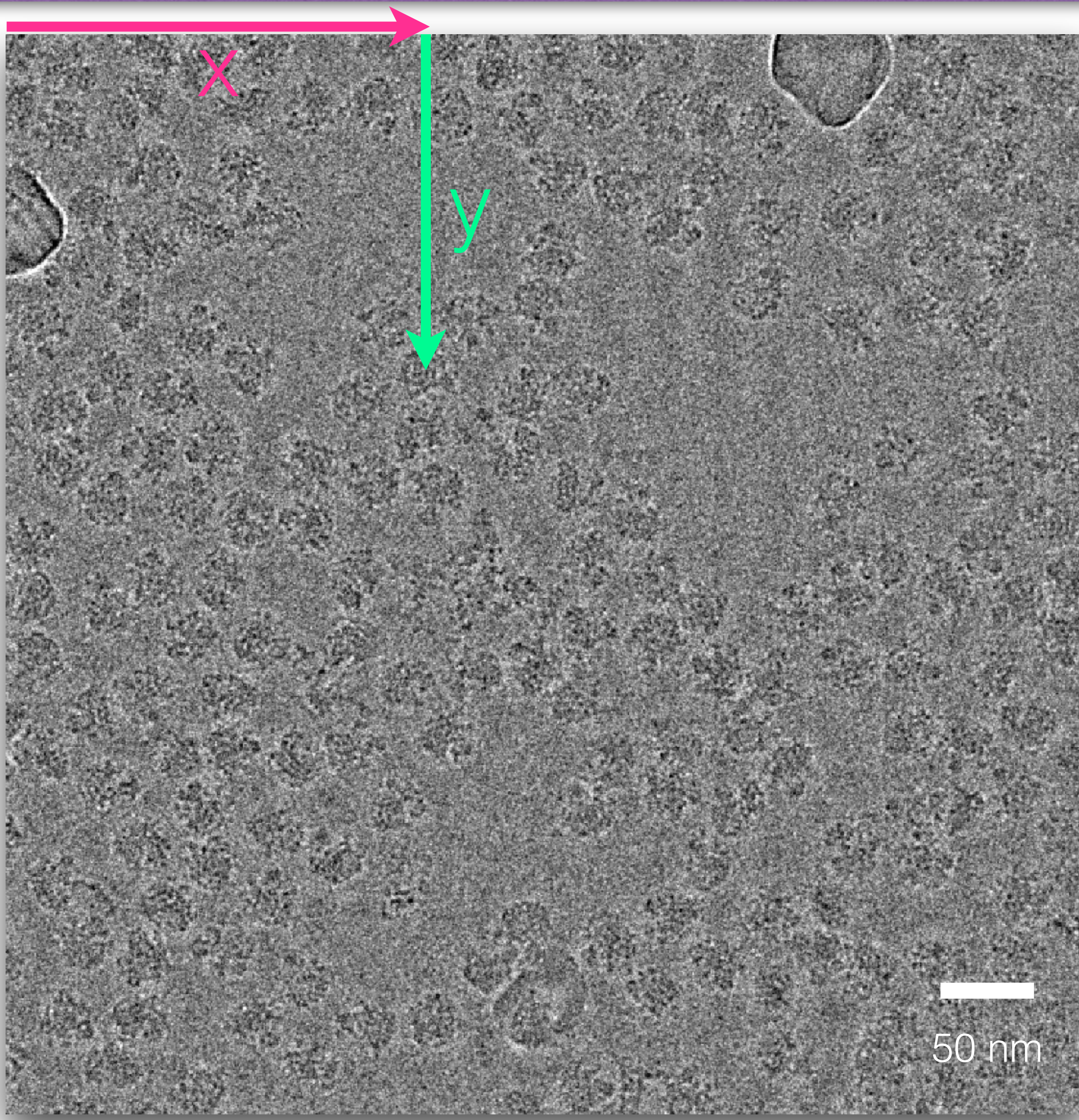


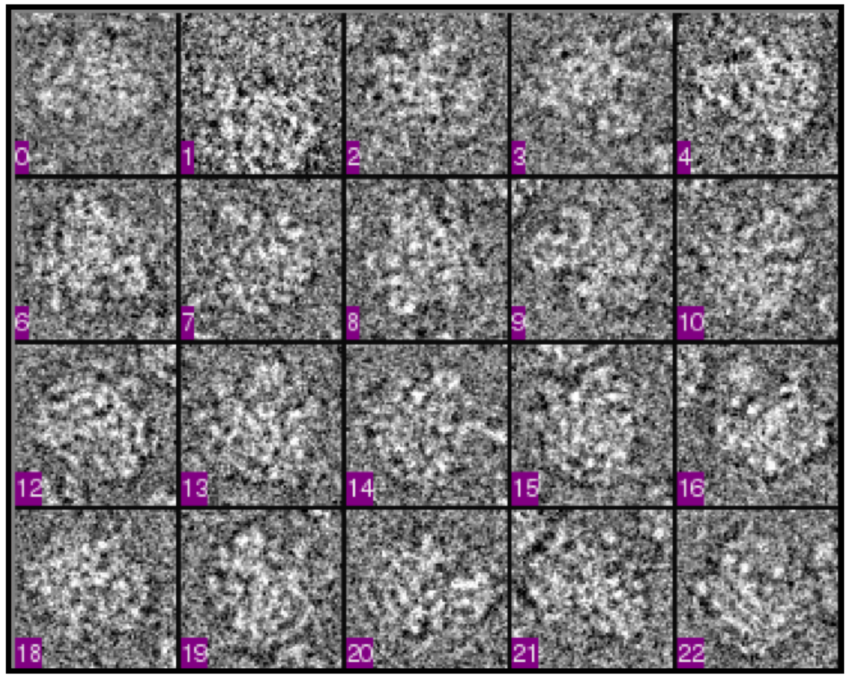
Unprocessed | Contrast flipped | Wiener filtered

$$\text{Wiener filtered Fourier image} = \text{Input Fourier image} \bullet \frac{\text{CTF}}{\text{CTF}^2 + \text{SNR}^{-1}}$$

Correction of Fourier amplitudes allows averaging

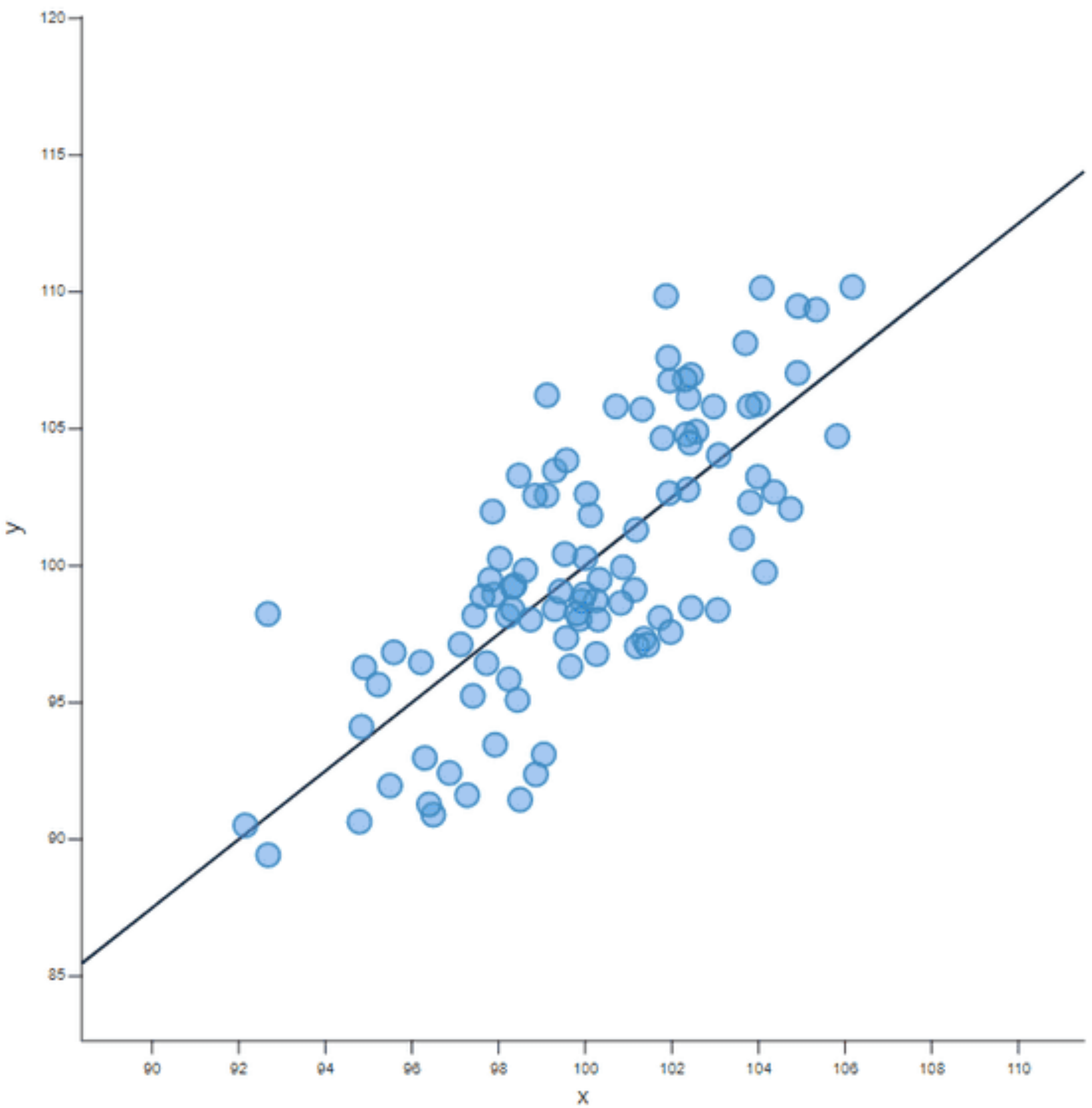
Centre of rotation - picking



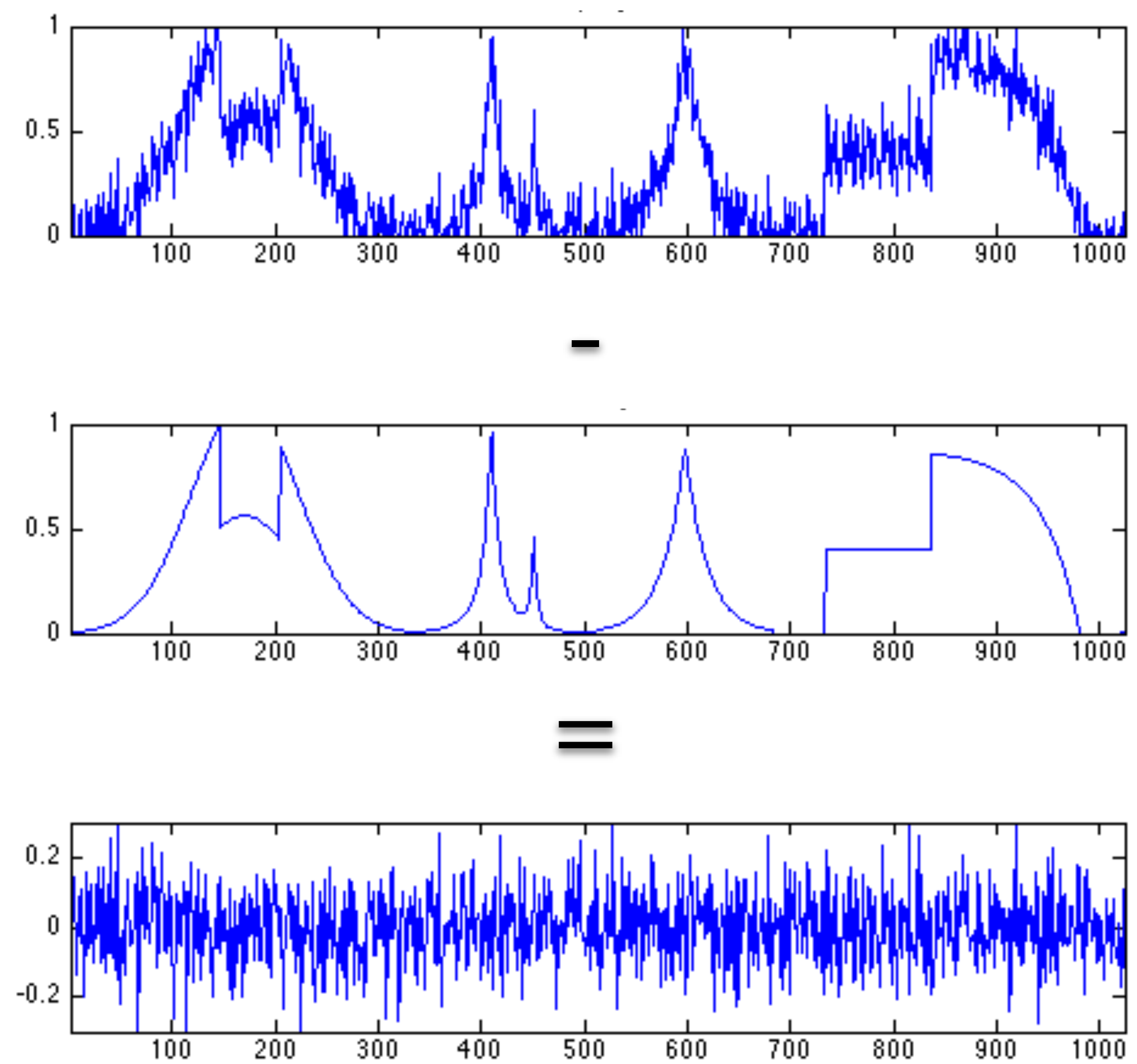


Two-steps to simplify process: picking / refinement

Quantifying image similarity



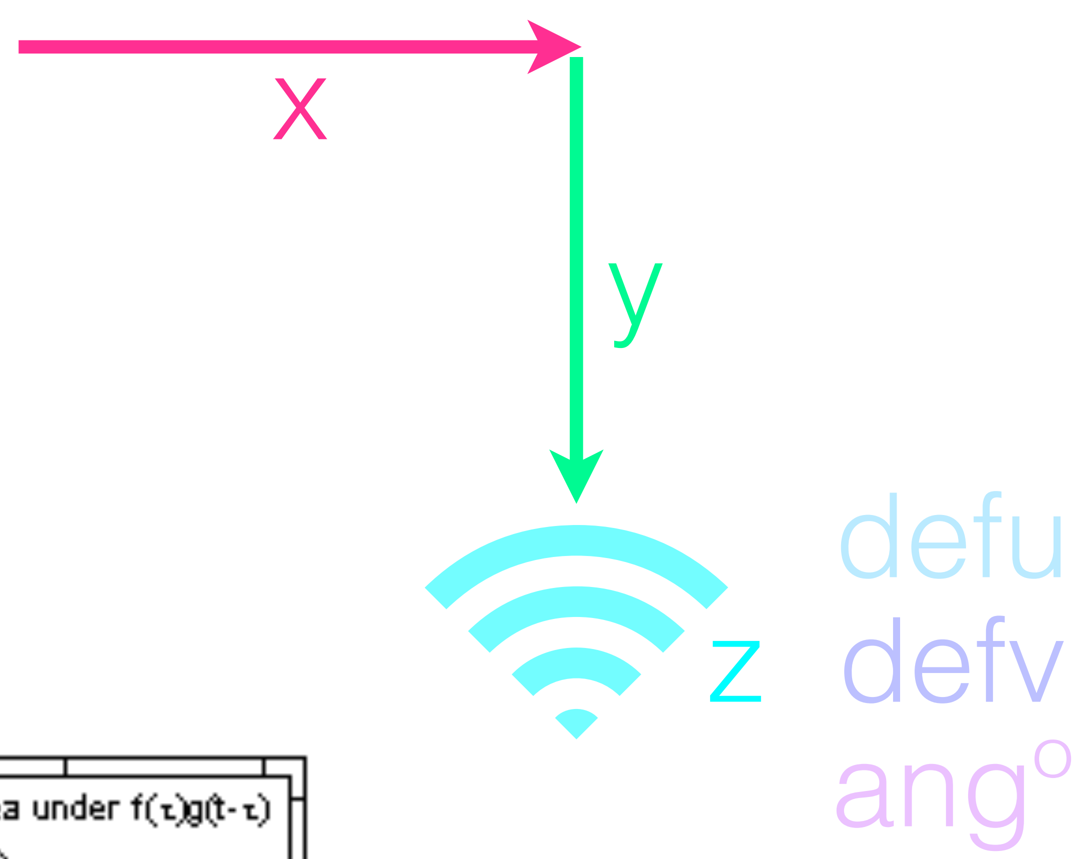
Correlation

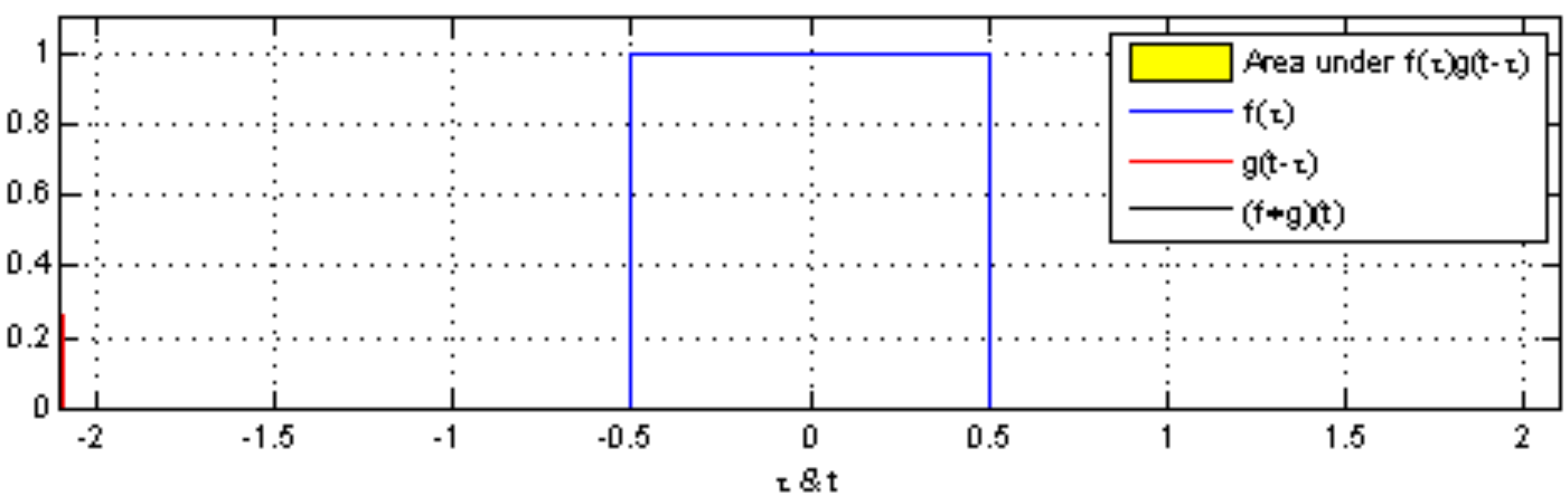


Likelihood

Correlation predominantly replaced by likelihoods

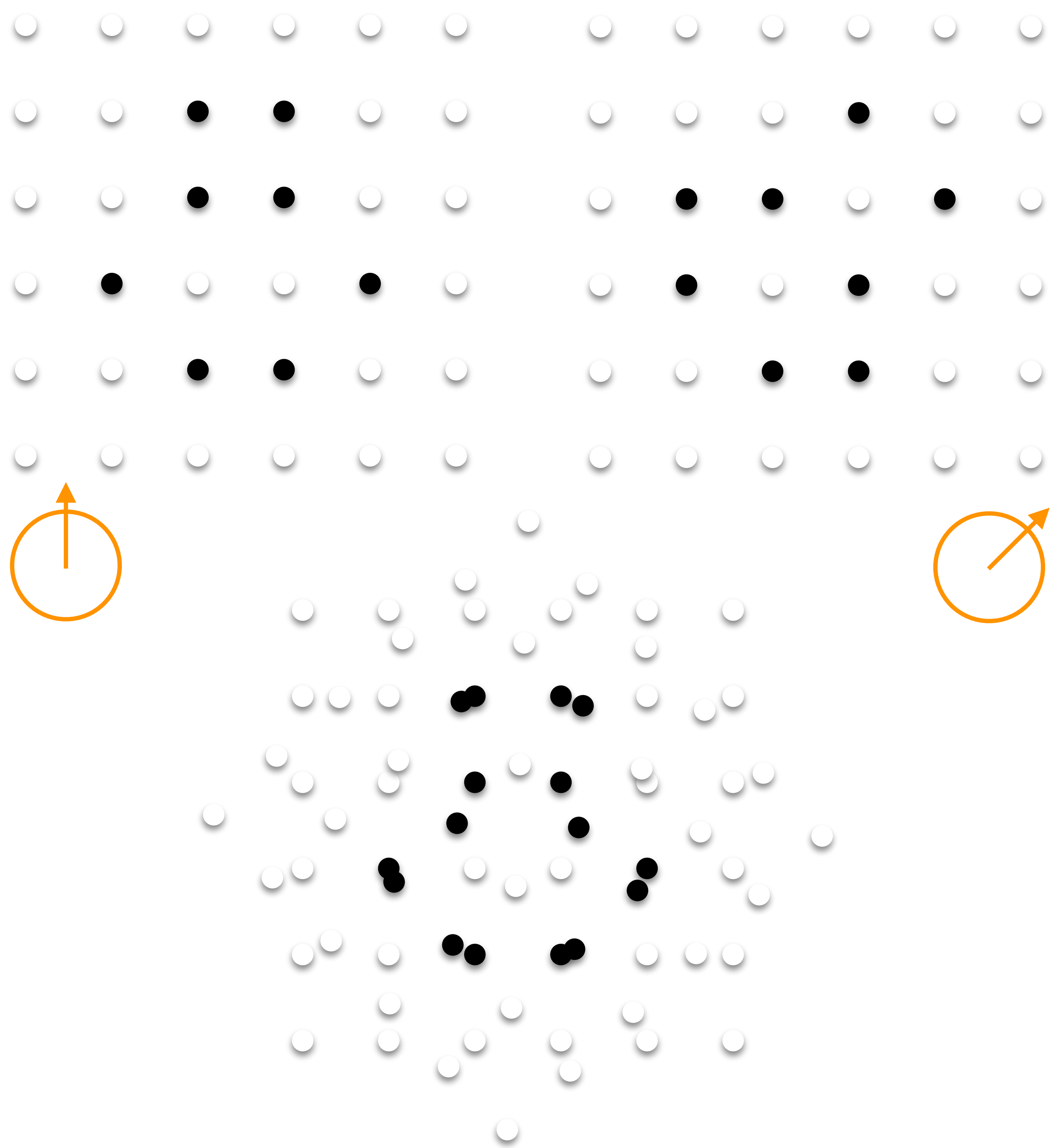
Alignment - x / y (+ frames), z (CTF)

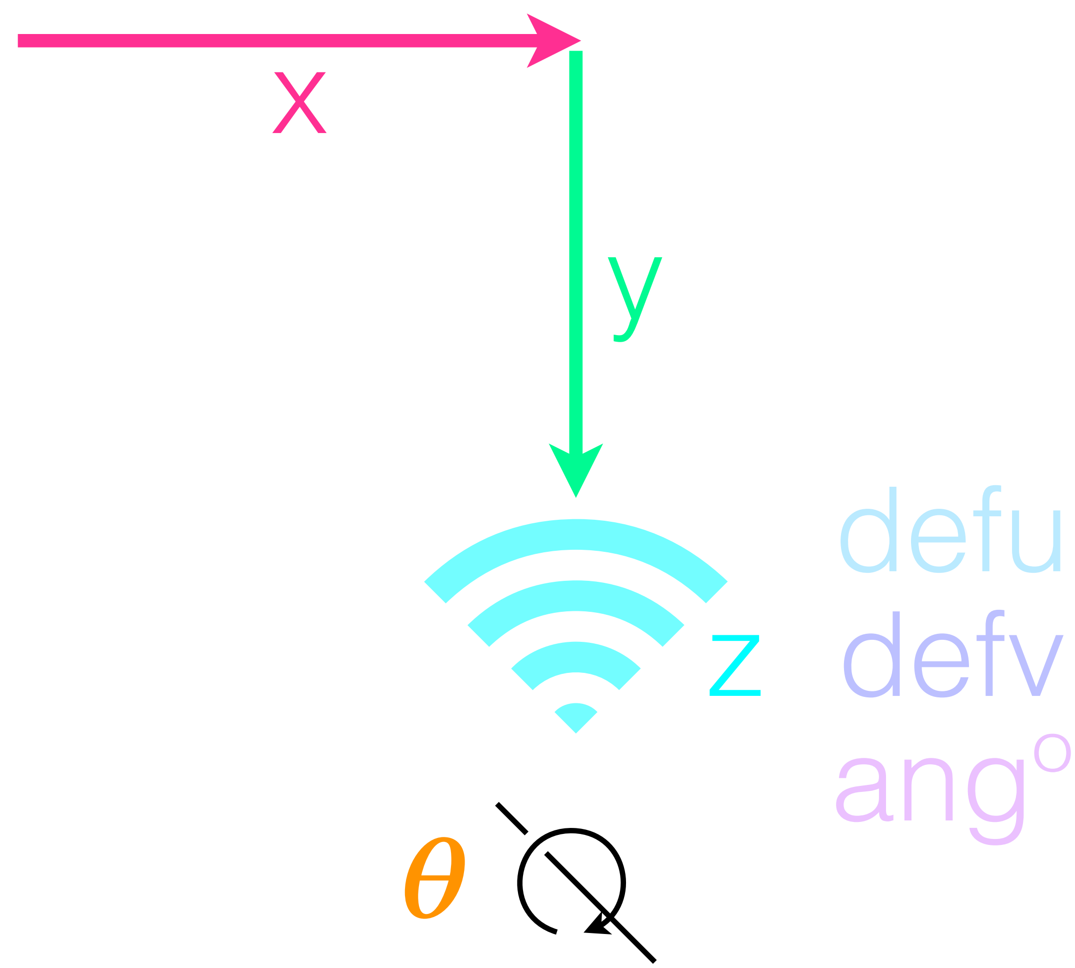




Entire process defined by a few parameters

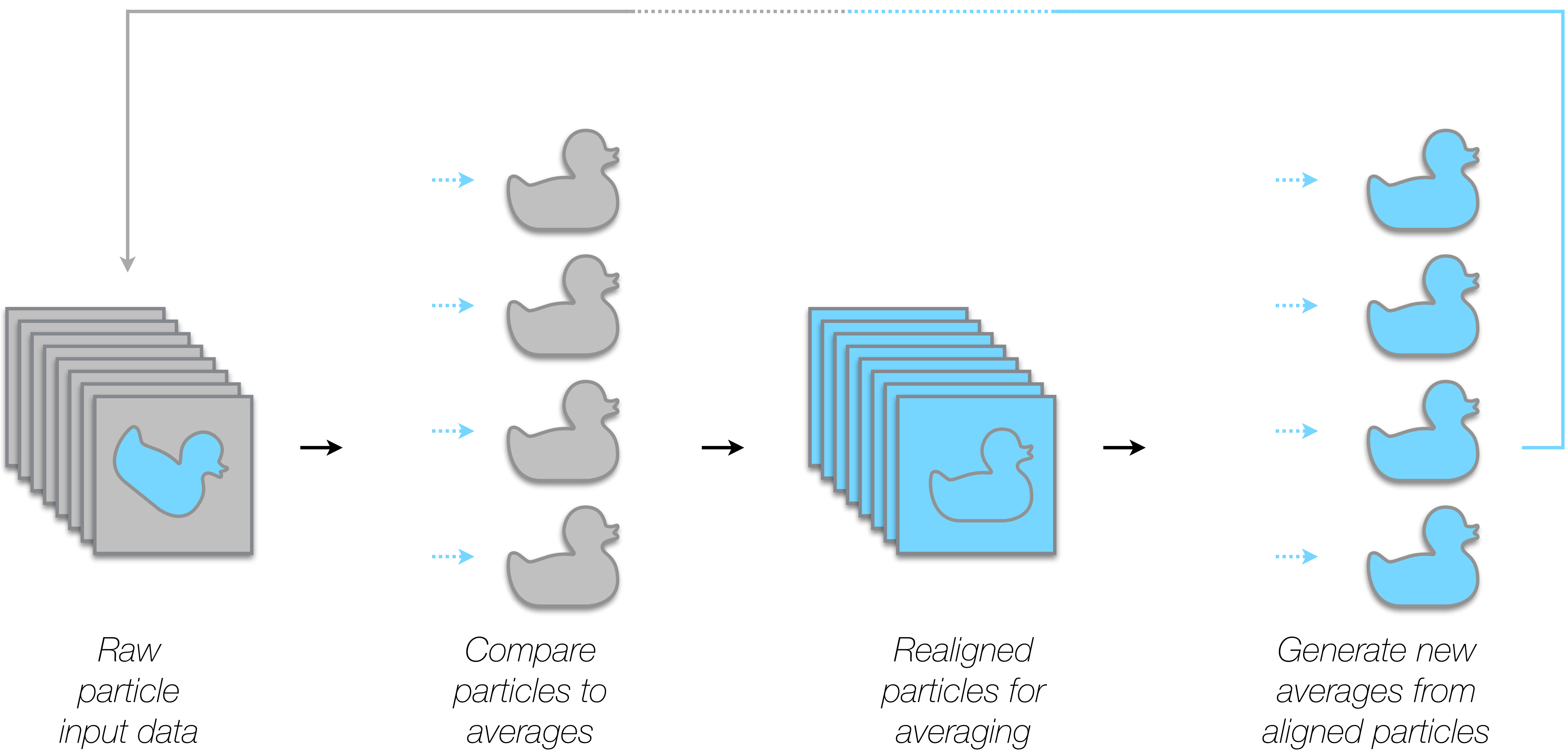
Alignment - angle of rotation





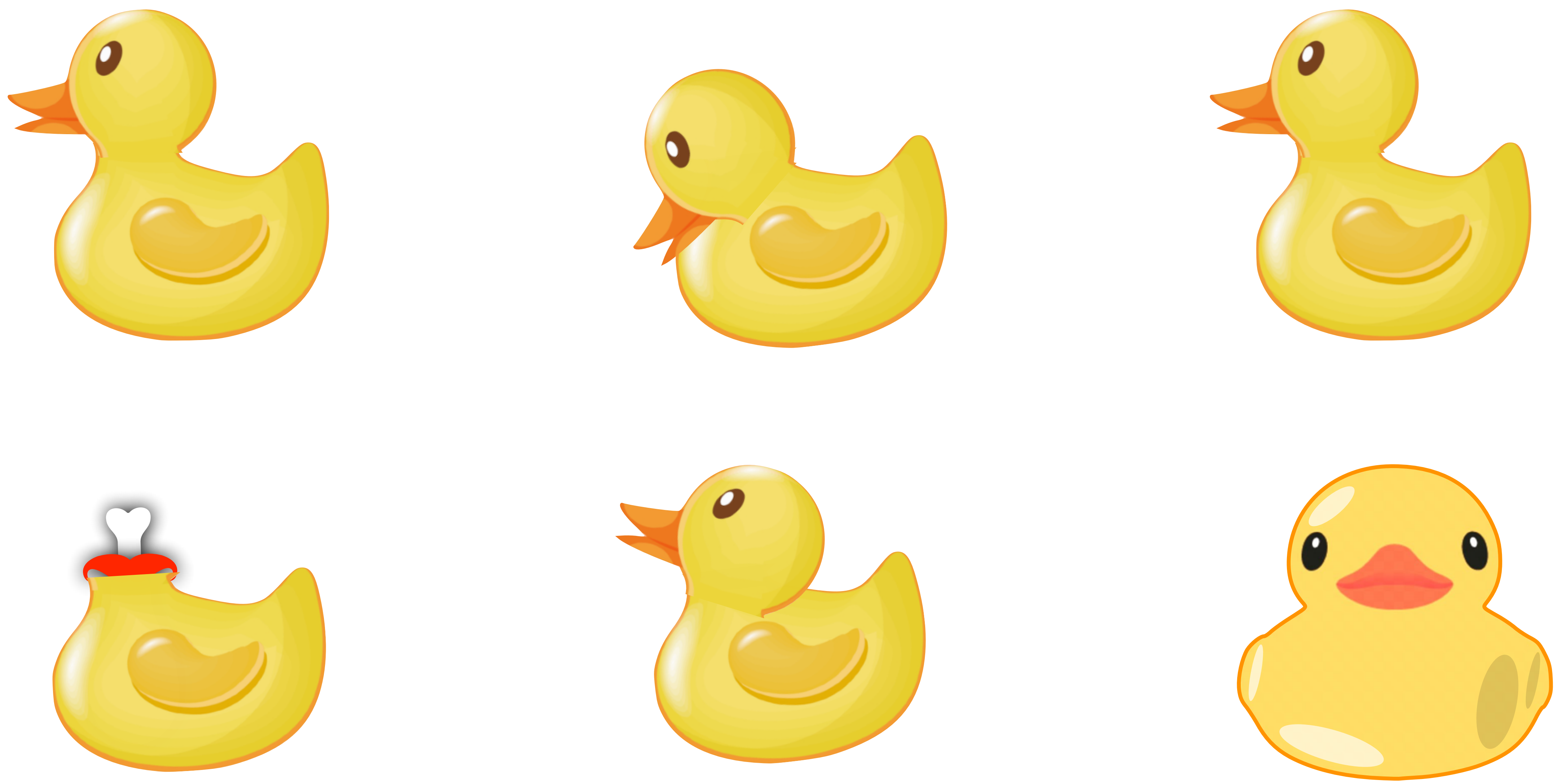
Lossy interpolation required for rotation

Iterative refinement to yield the aligned average



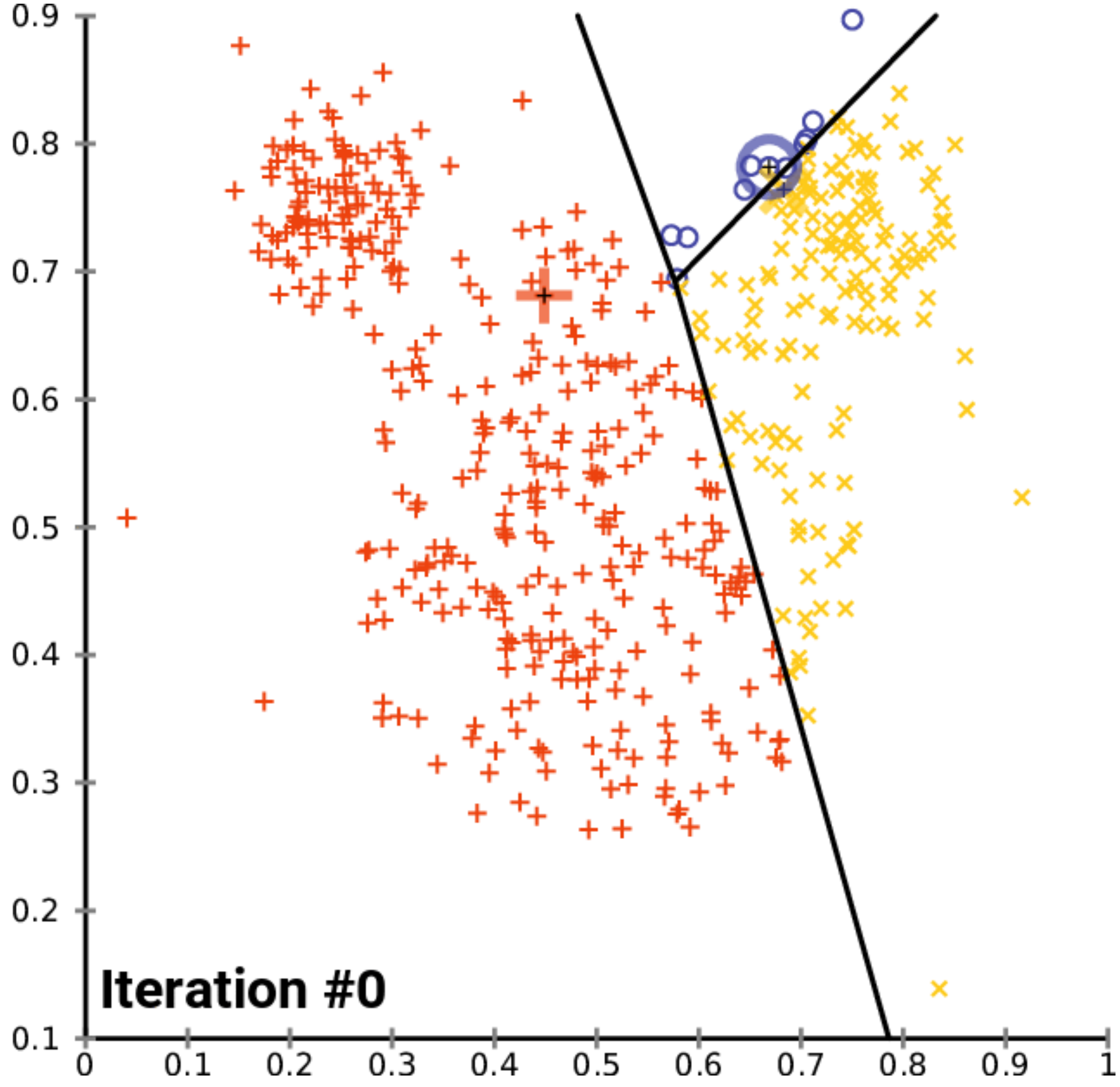
Well-behaved process which generally converges

Heterogeneity - causes



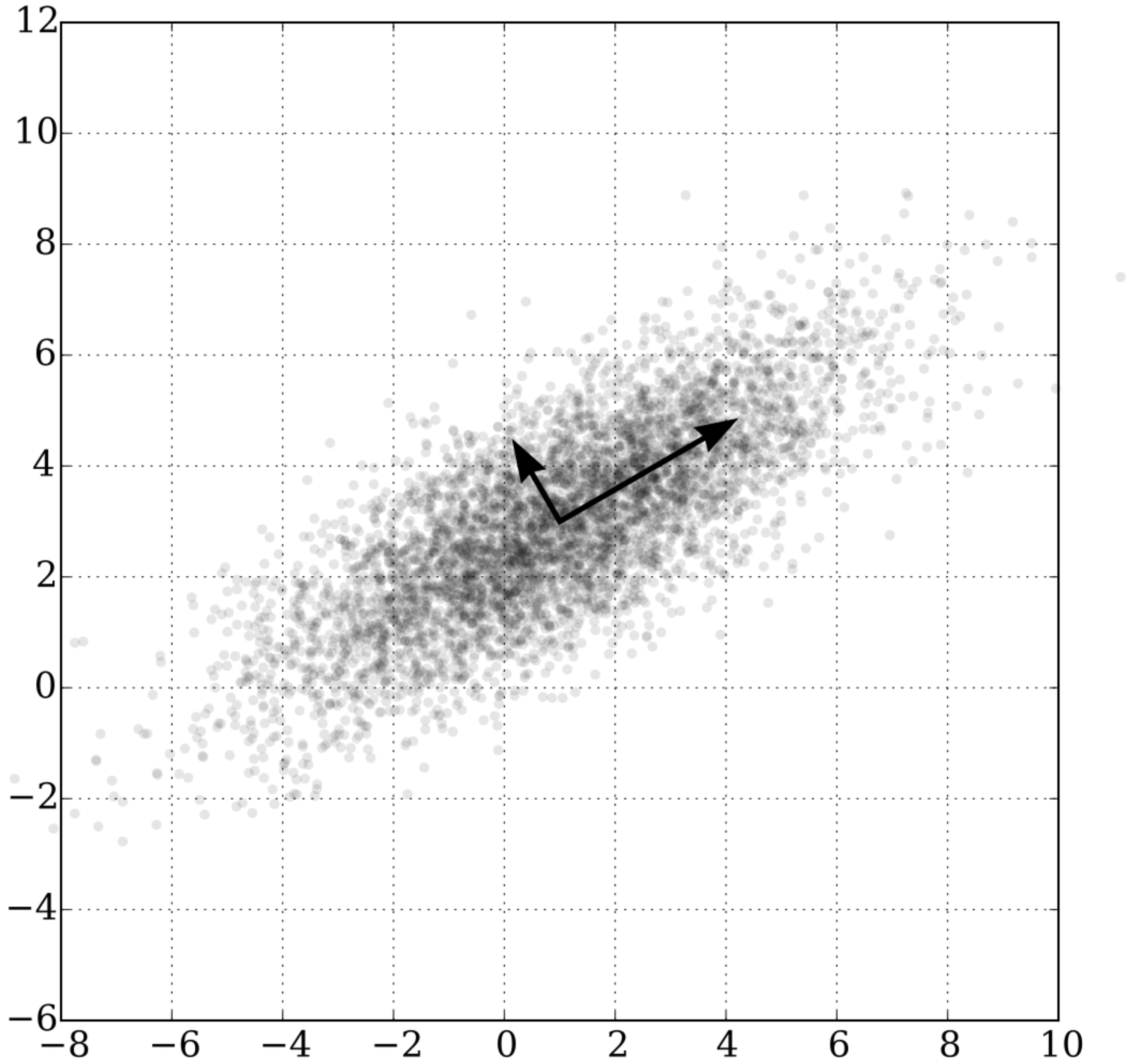
Compositional and conformational vs. angular

Heterogeneity - K-means classification



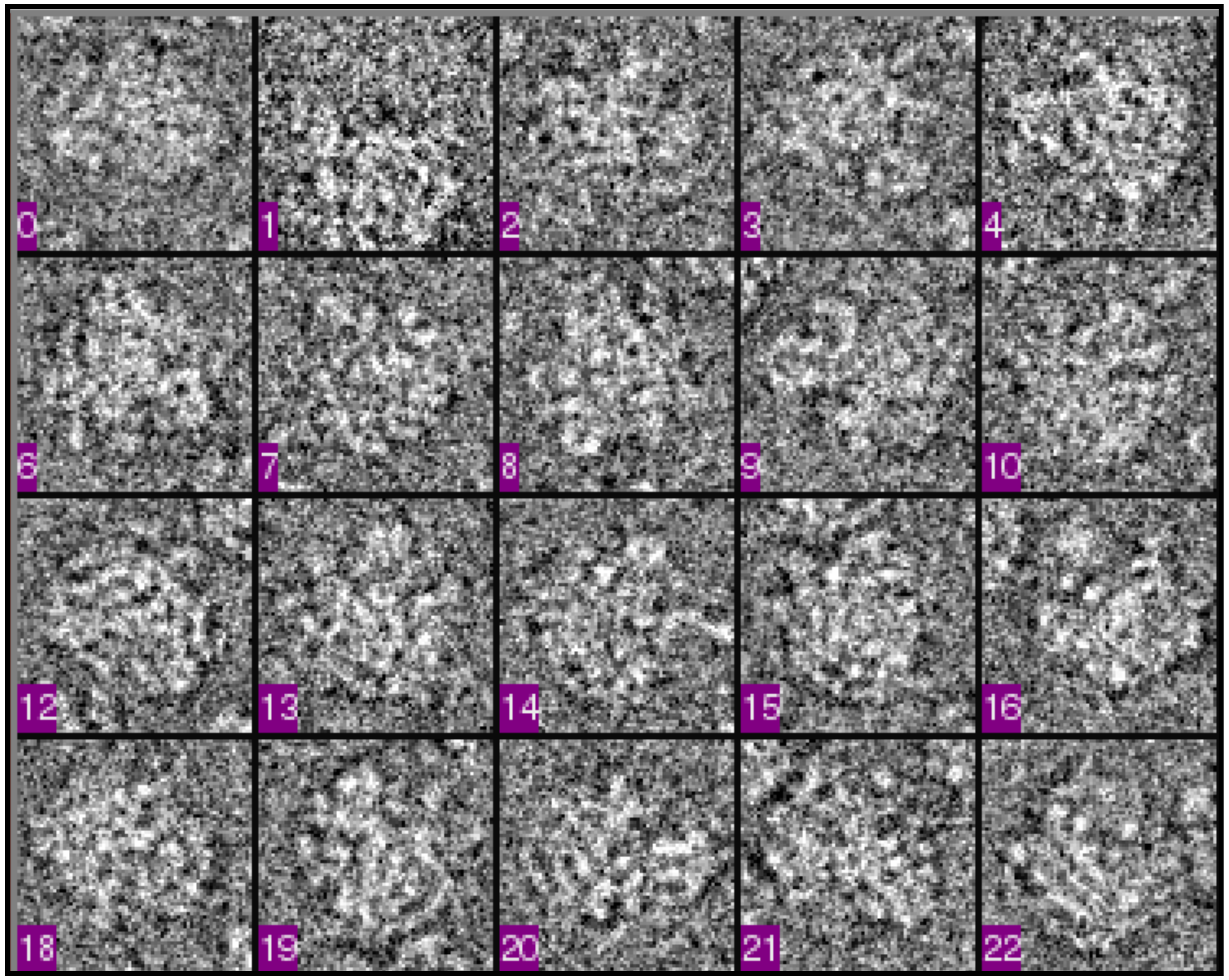
Extensible family of emergent classifiers

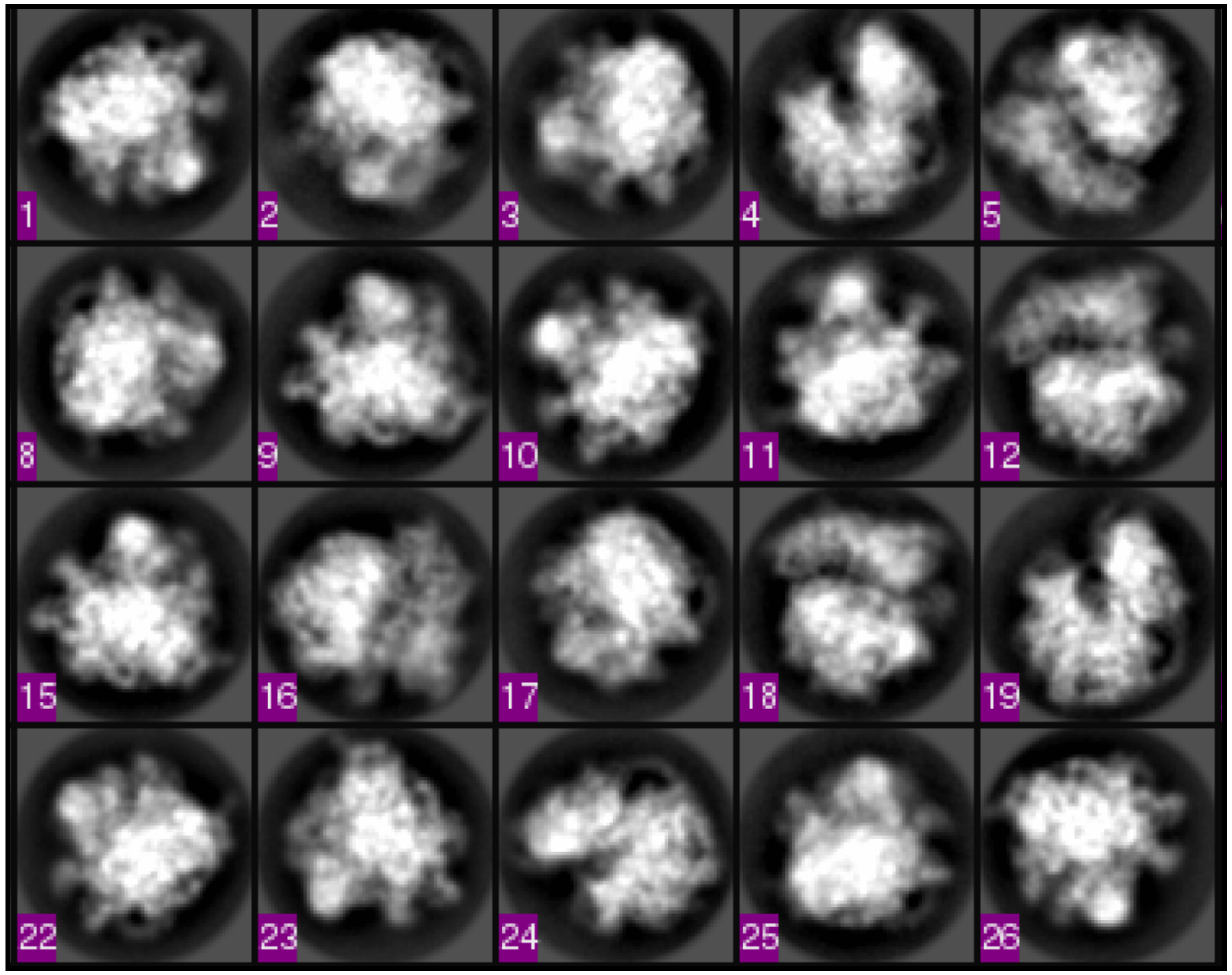
Heterogeneity - principal component analysis



More complex but more powerful classification

2D averaging - model / results



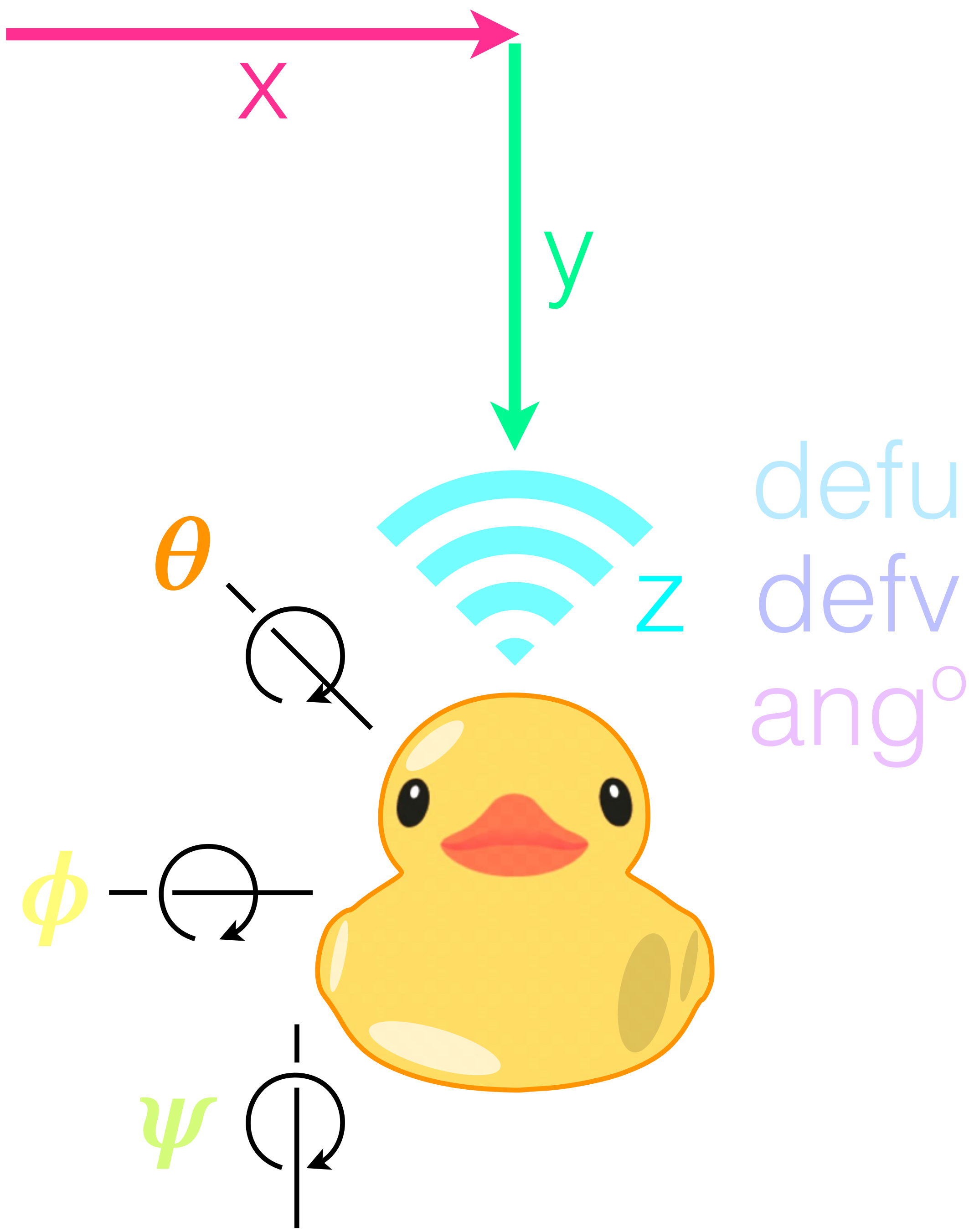


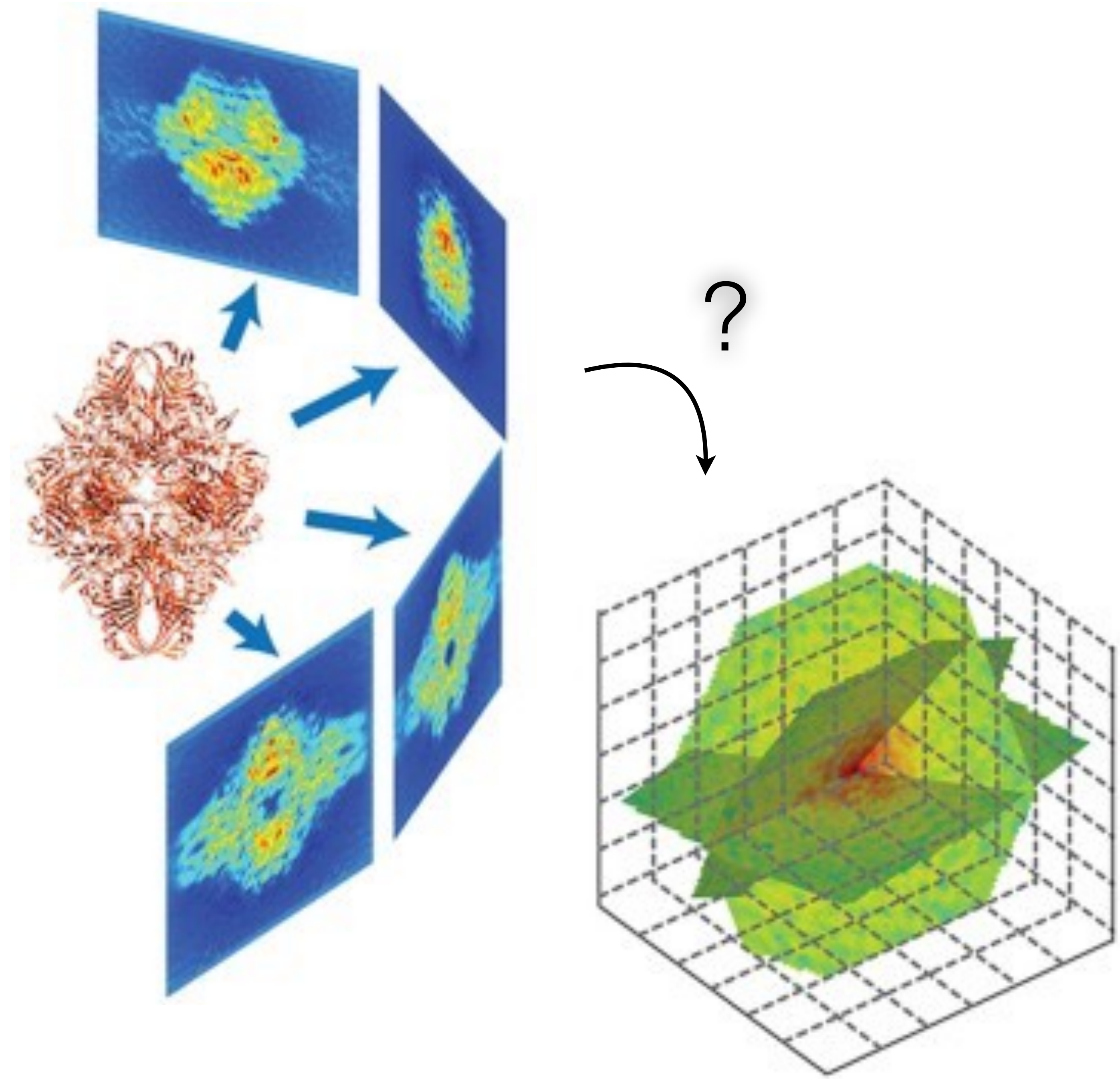
High resolution averages of preferred orientations

3D reconstruction - the problem

- classically “ill-posed” and requires regularisation

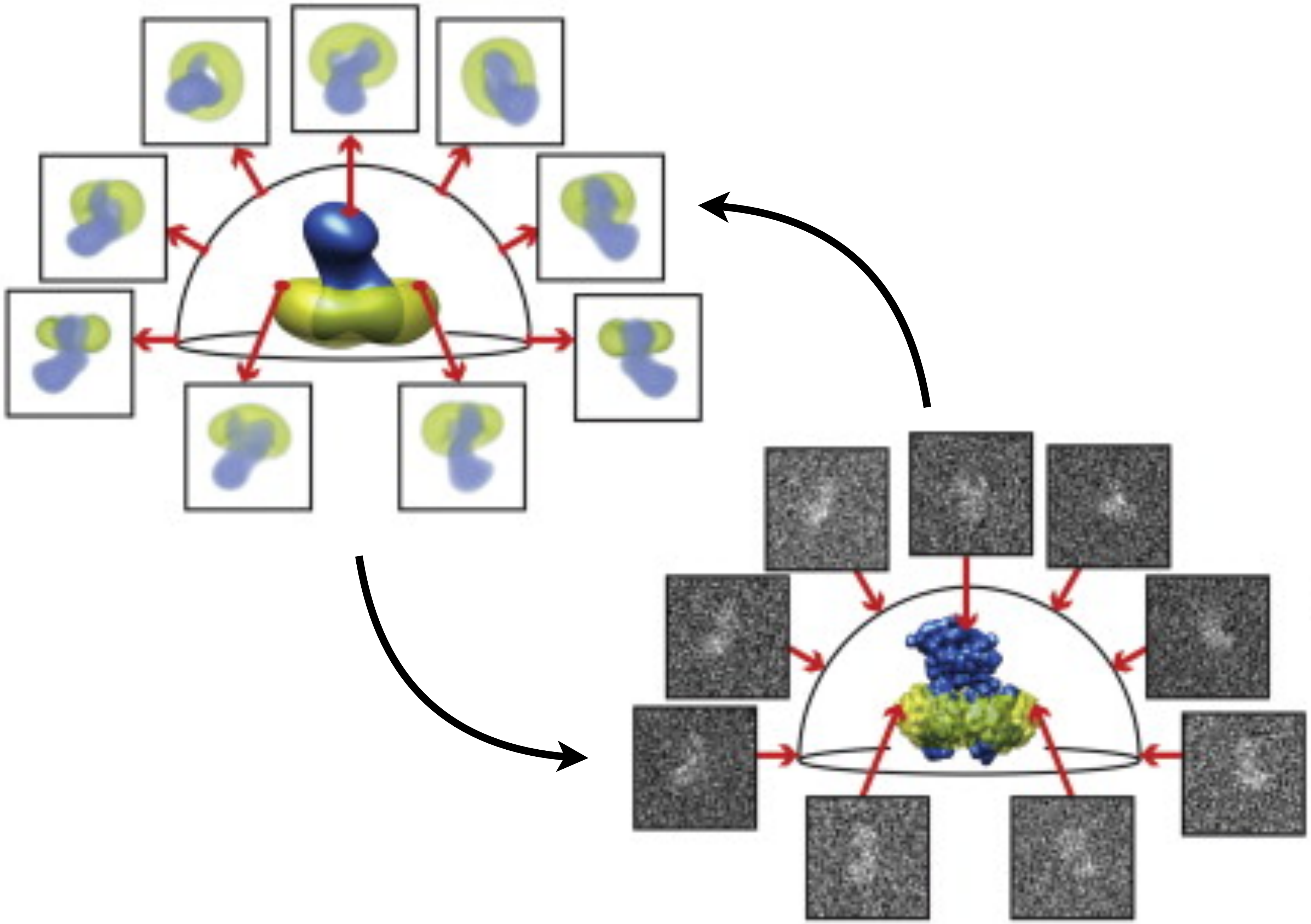
3D reconstruction - regularisation





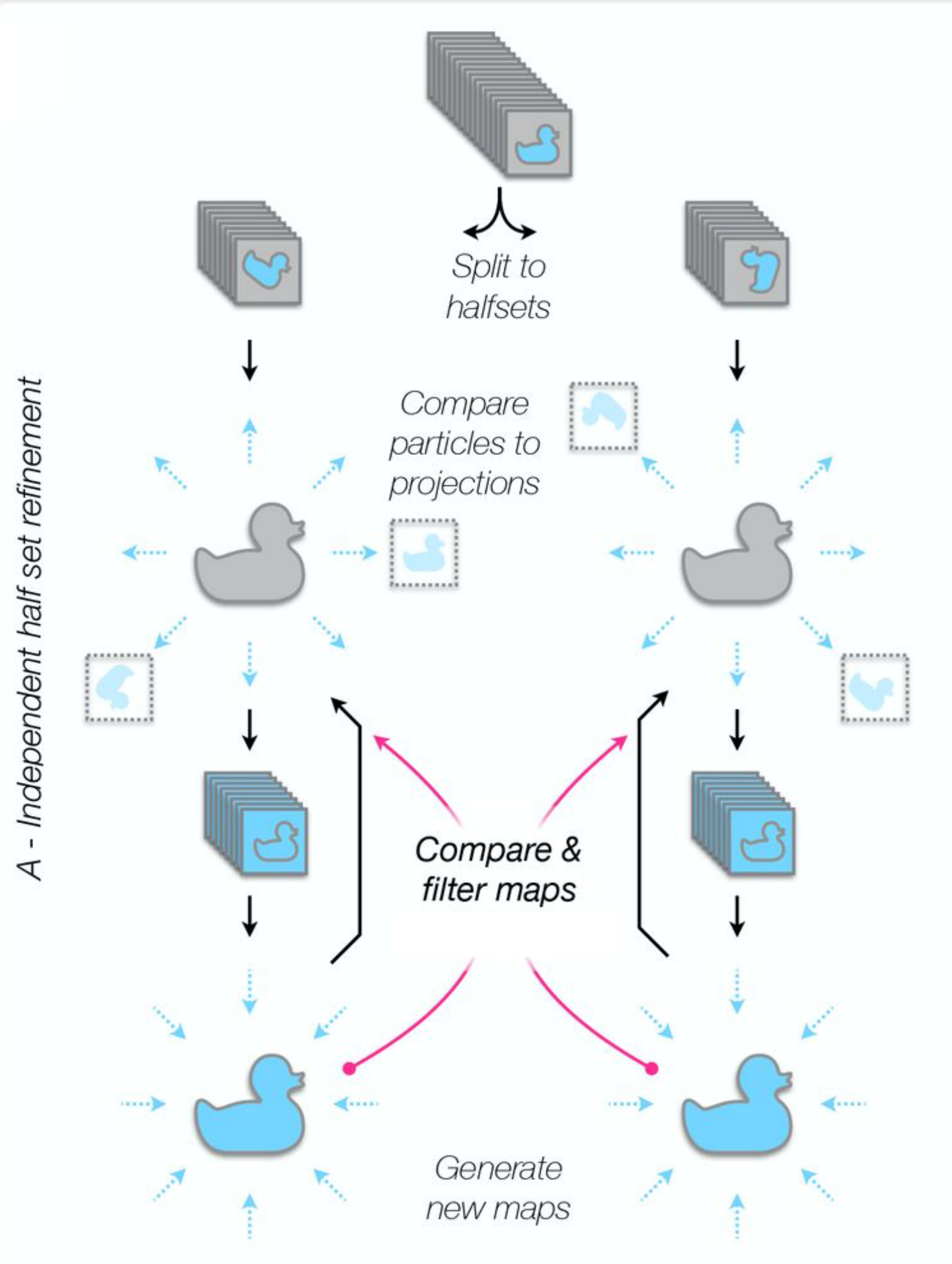
Assumptions - smooth / complete / initial volume

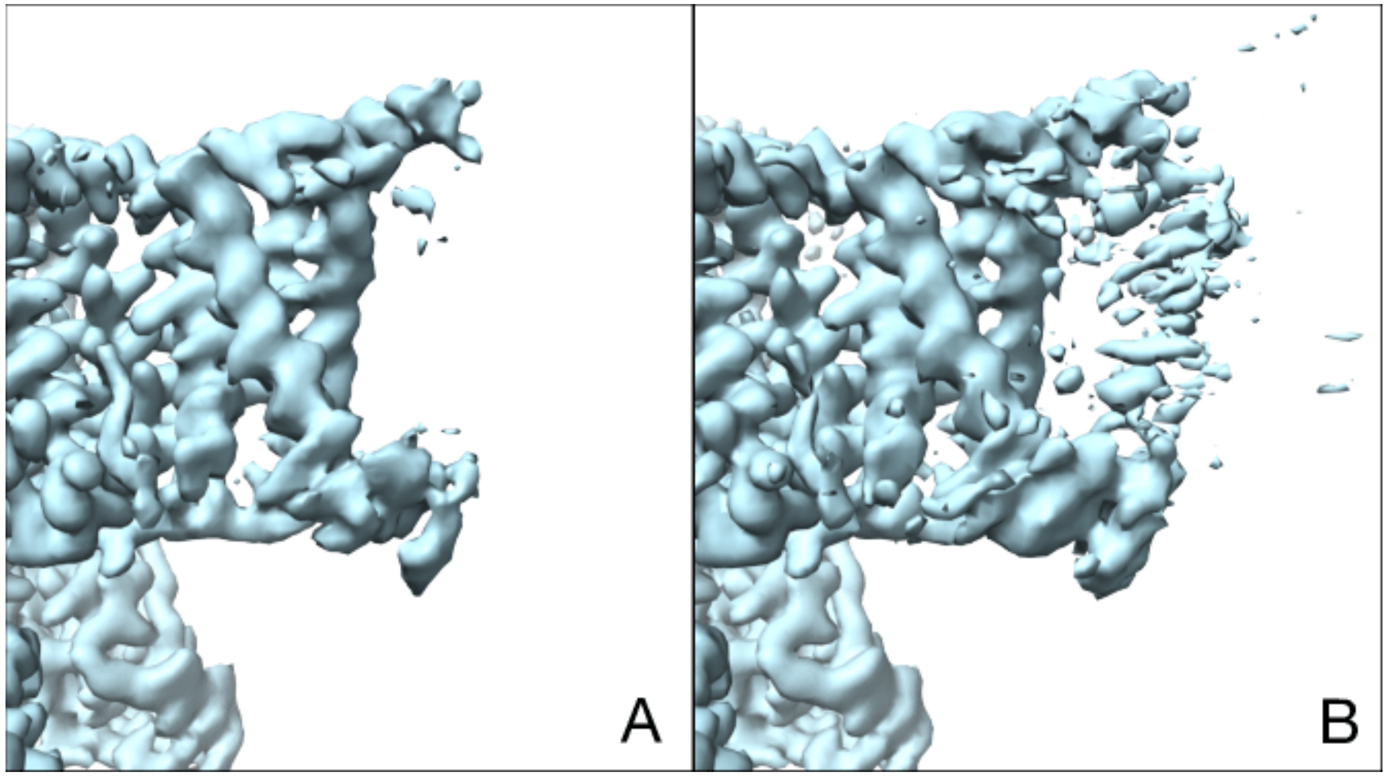
Iterative refinement in 3D



Two more angles - Initial volume required

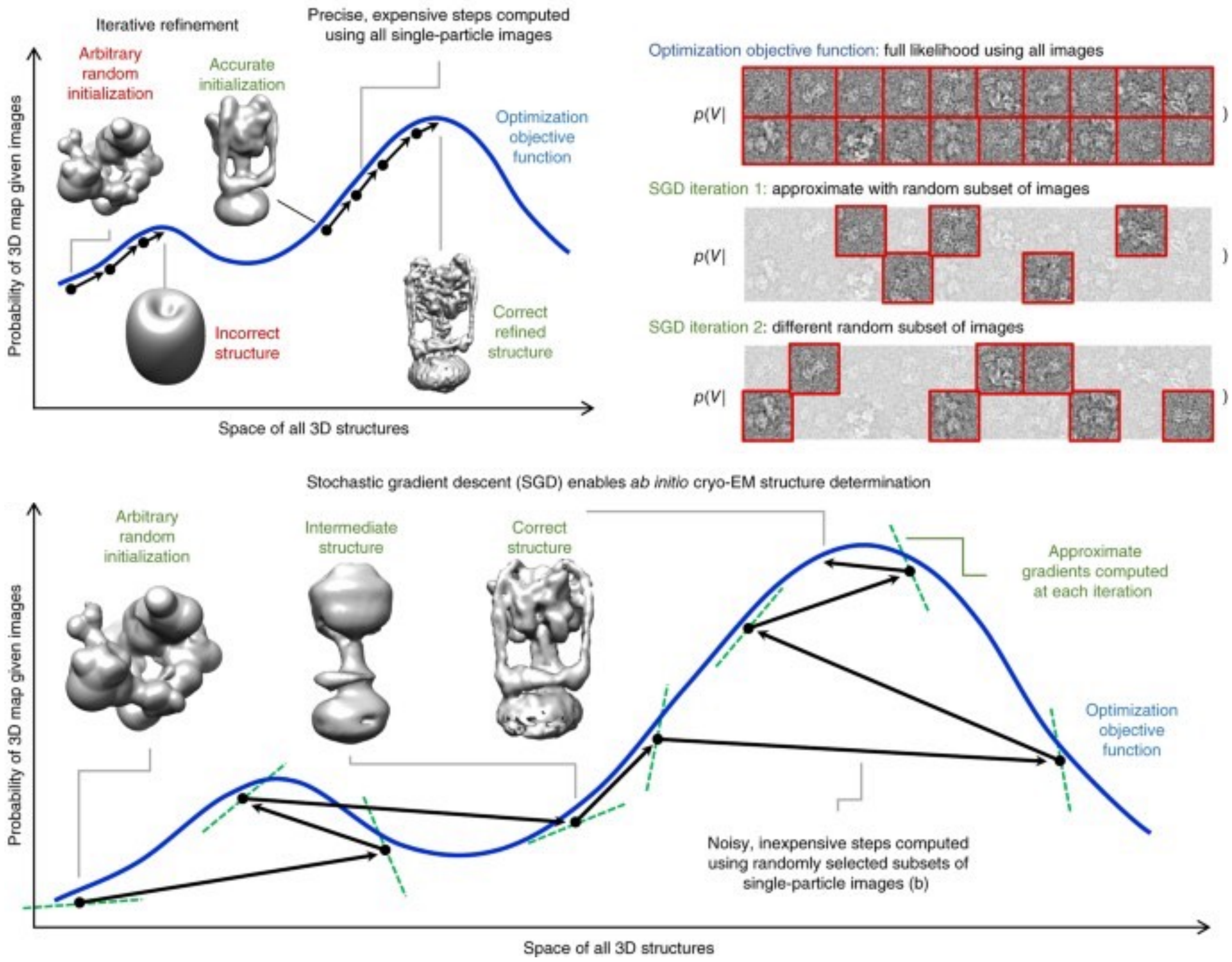
Overfitting and filtering - half sets





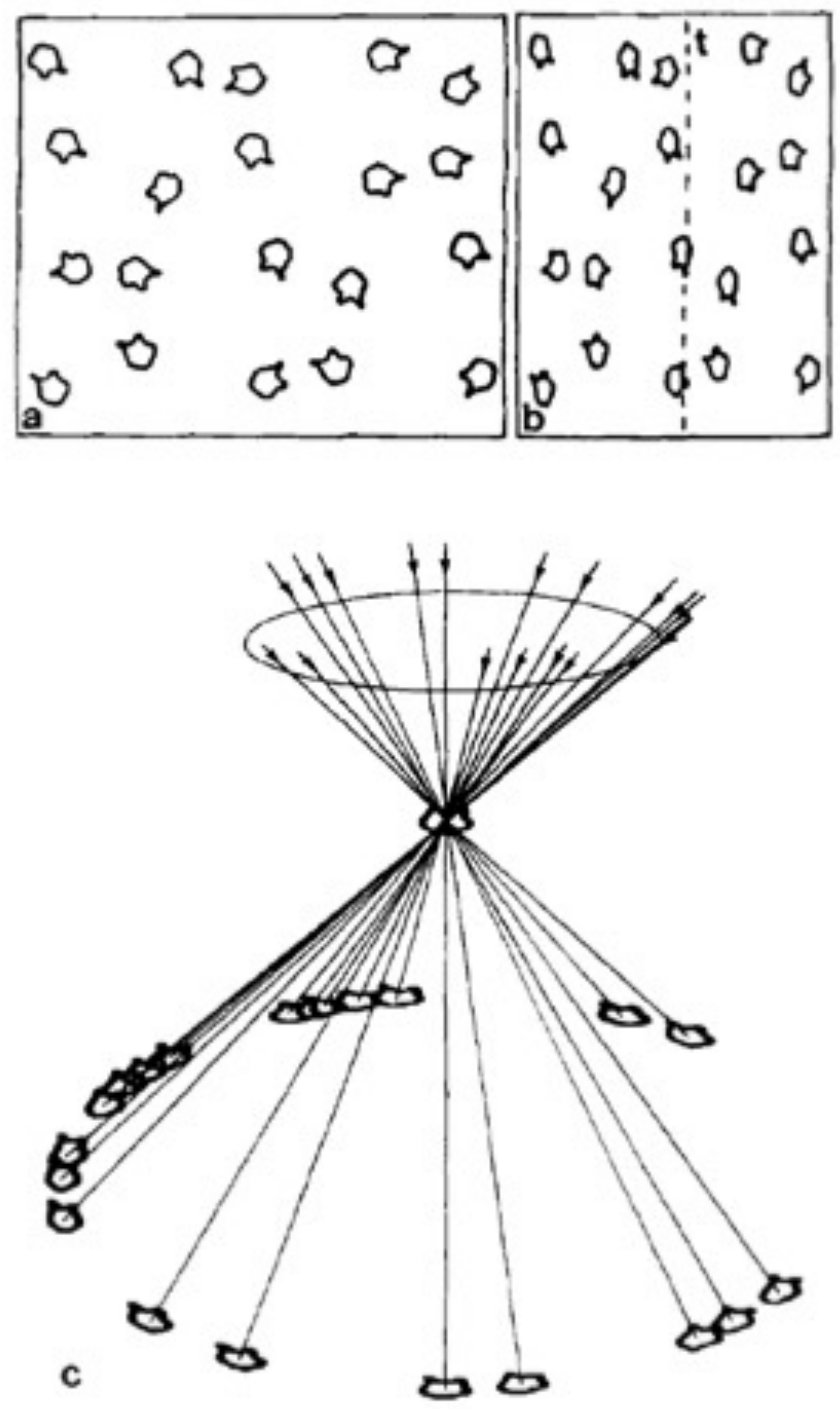
Full independence required to avoid overfitting

Initialisation - stochastic gradient descent



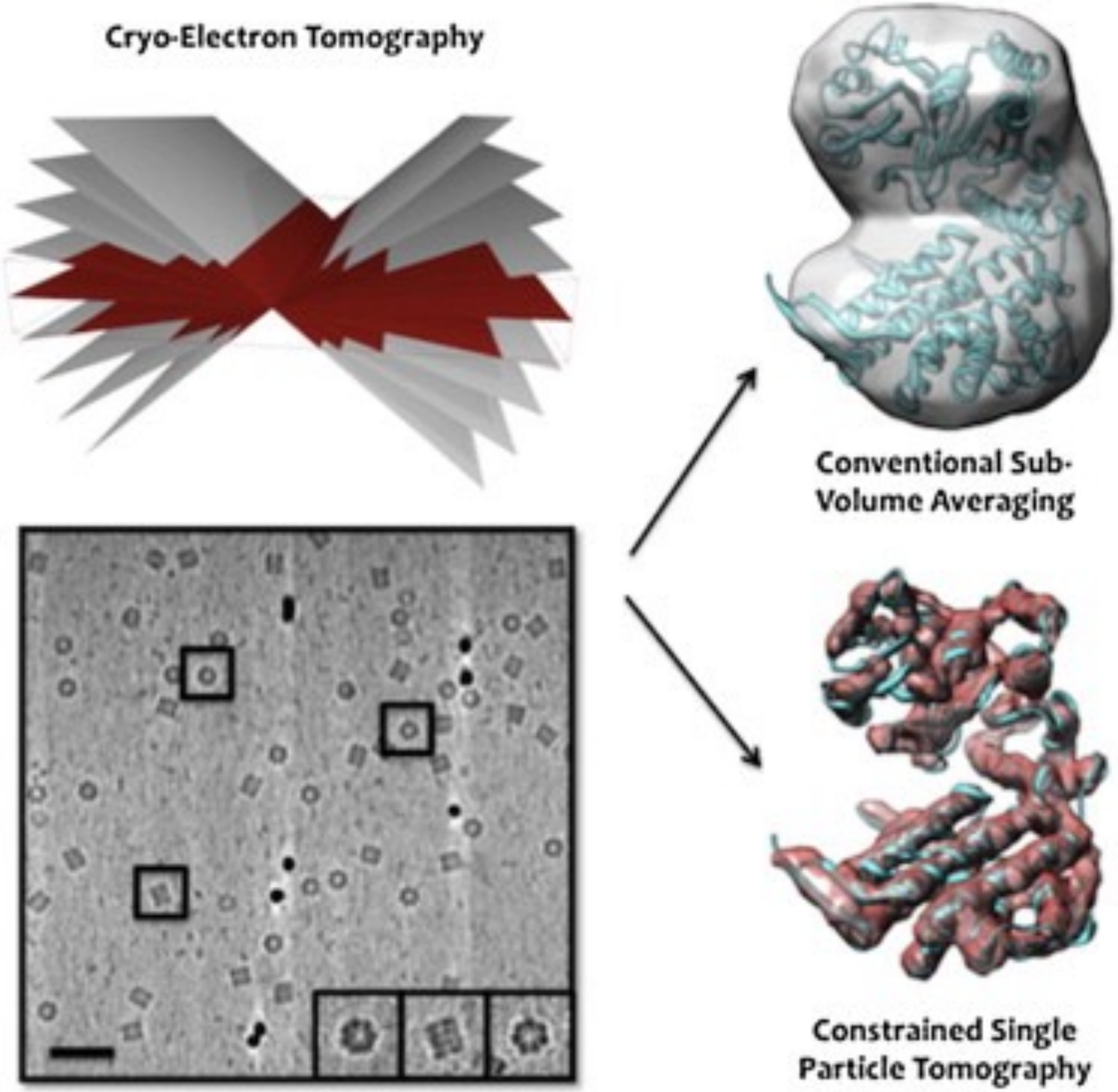
Like “Simulated annealing” - can be wrong / hand

Initialisation - random conical tilt



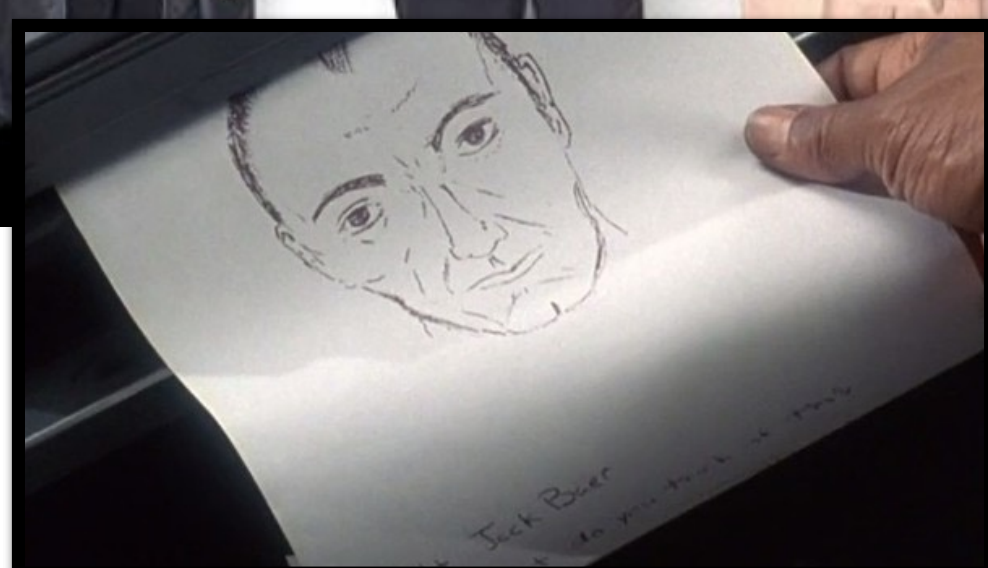
Assign angles from two different images of sample

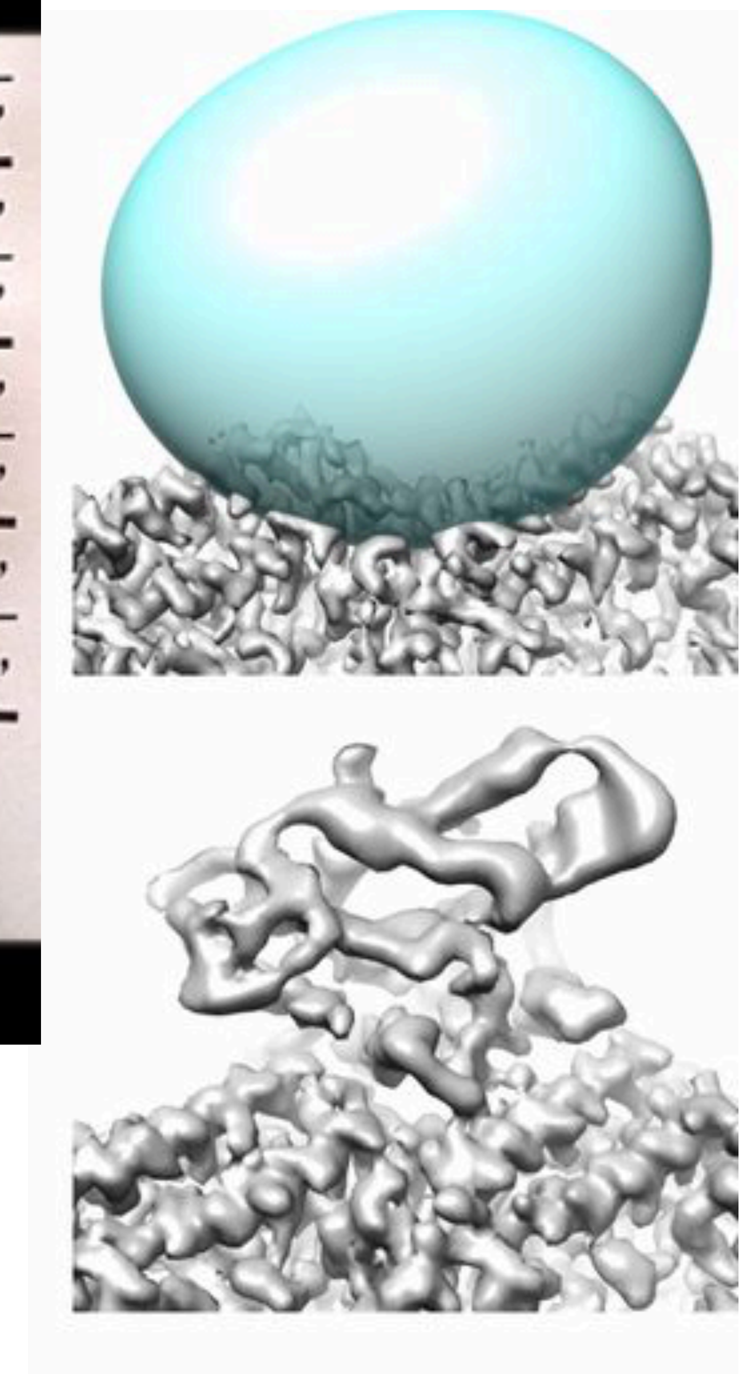
Initialisation - tomography



Best approach - no requirements or caveats

Heterogeneity - reference based classification

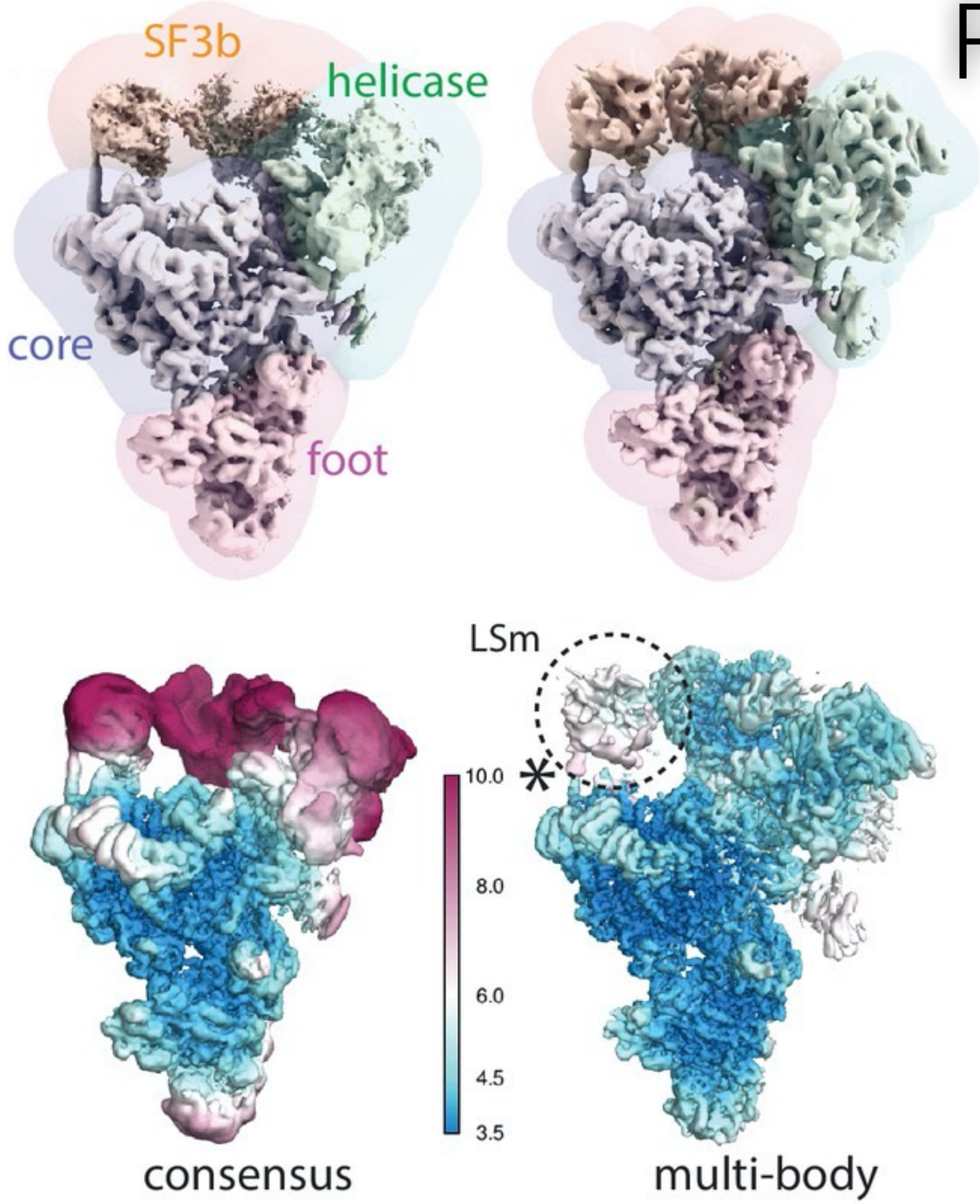


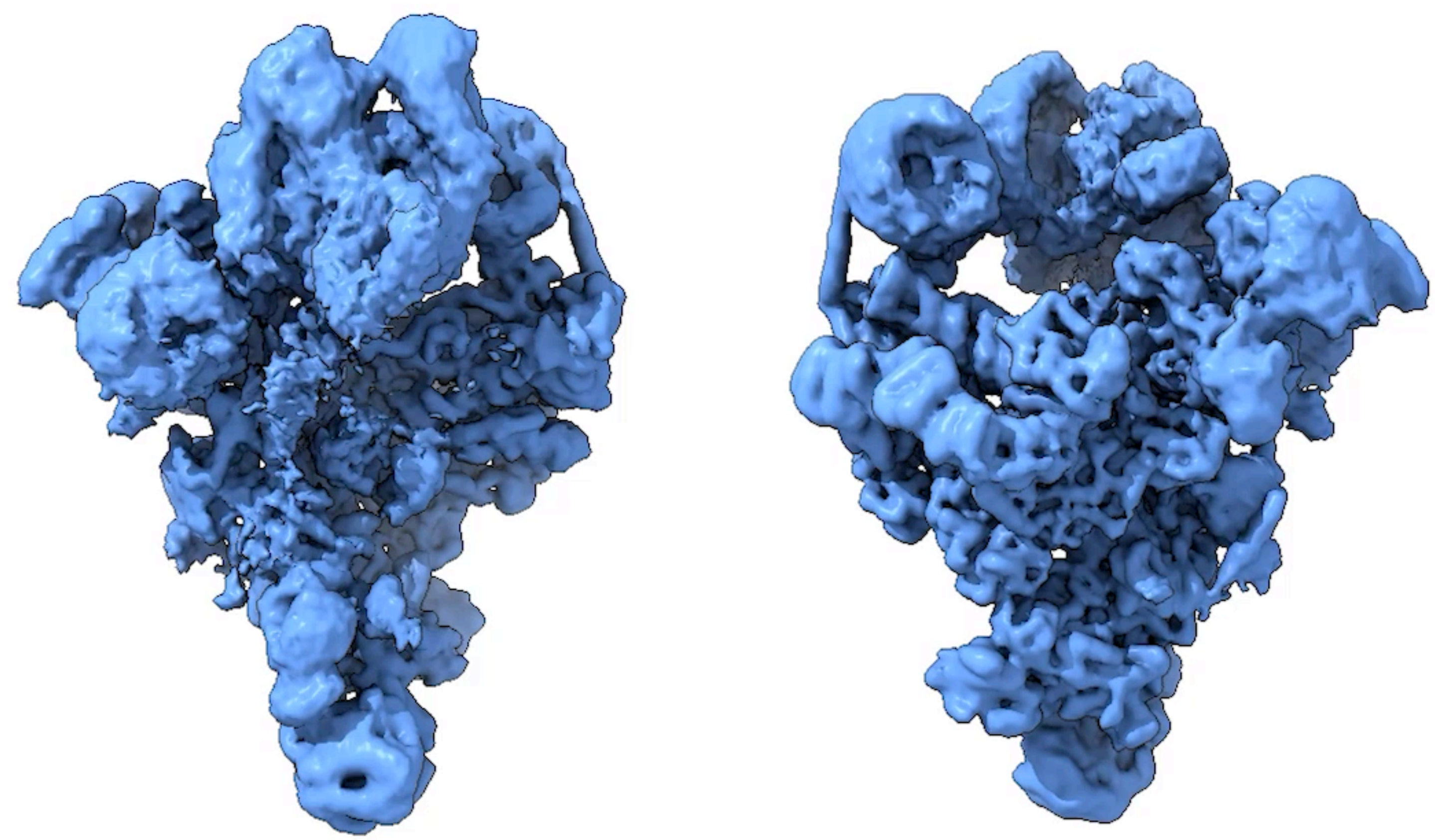


K-means family - masking / without alignment

Heterogeneity - multiple bodies vs. deep learning

RELION

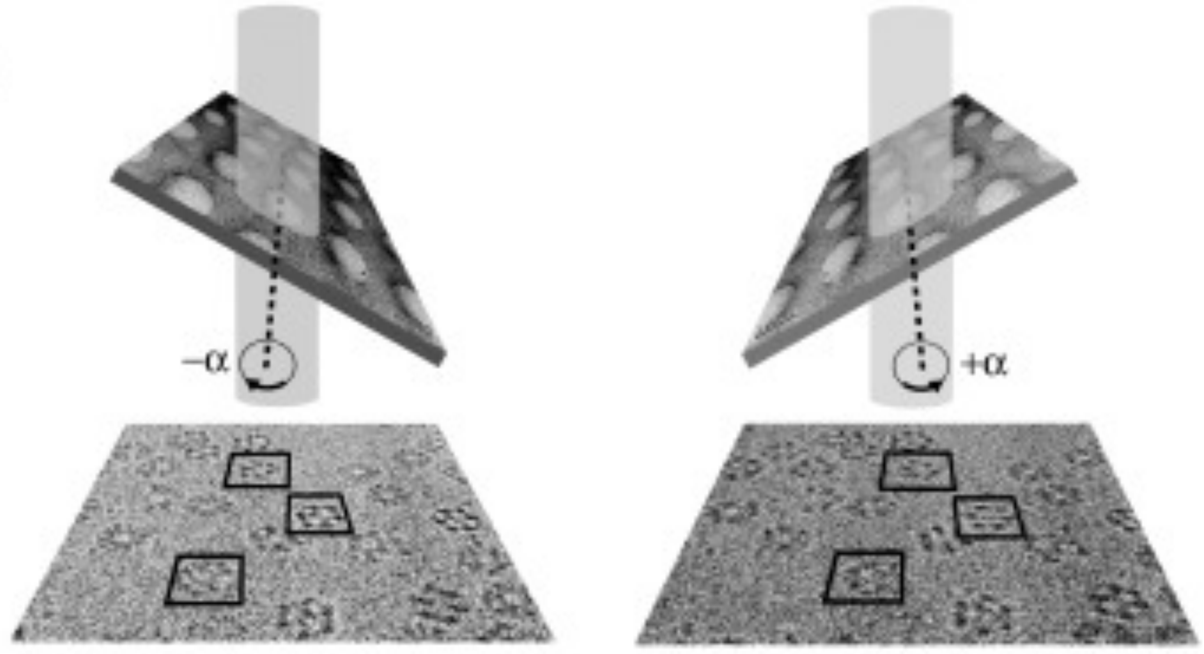


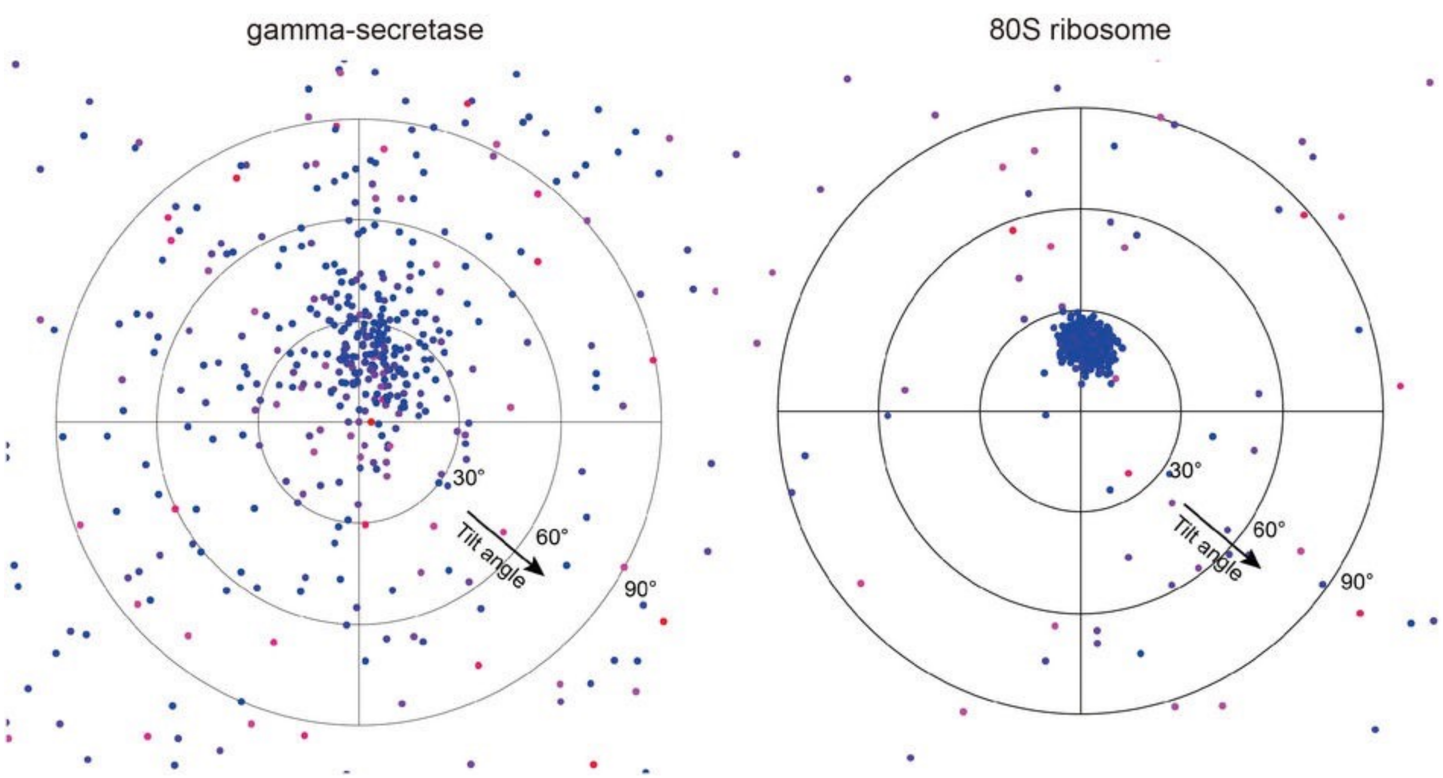


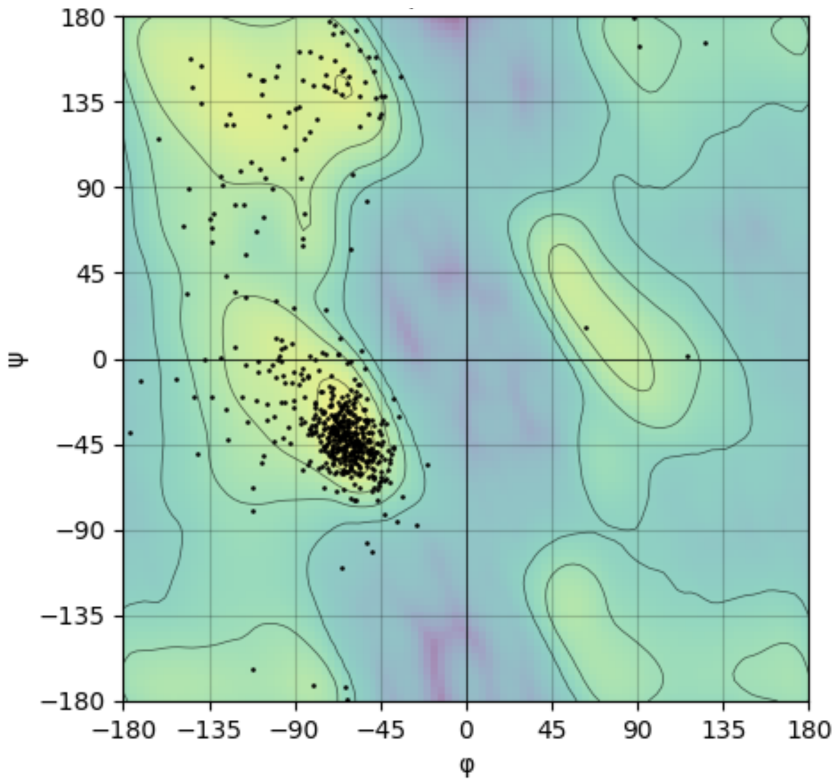
CryoDRGN

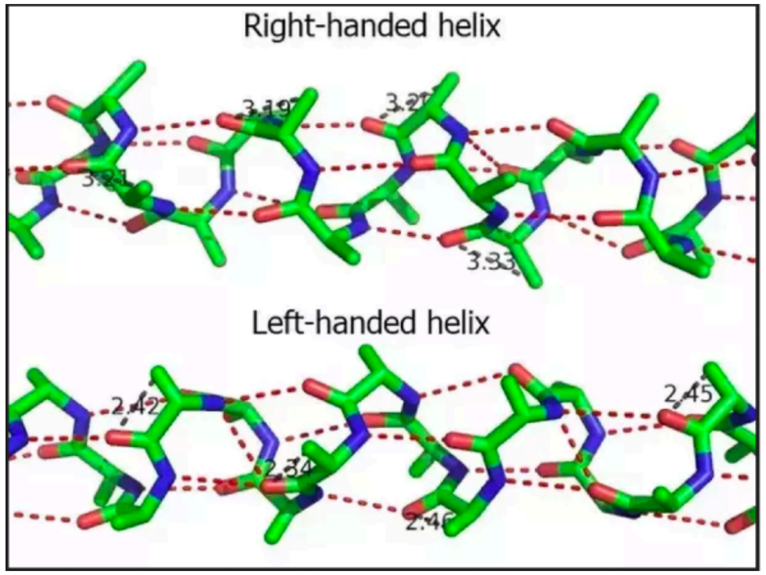
New techniques being tested and optimised

Hand





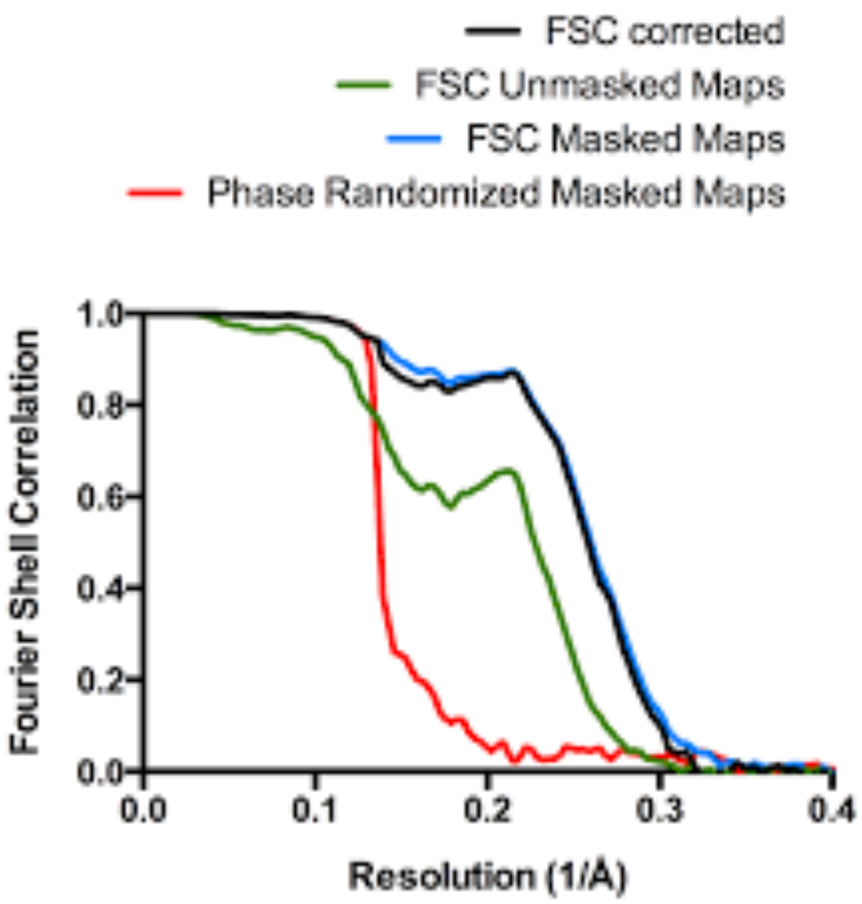


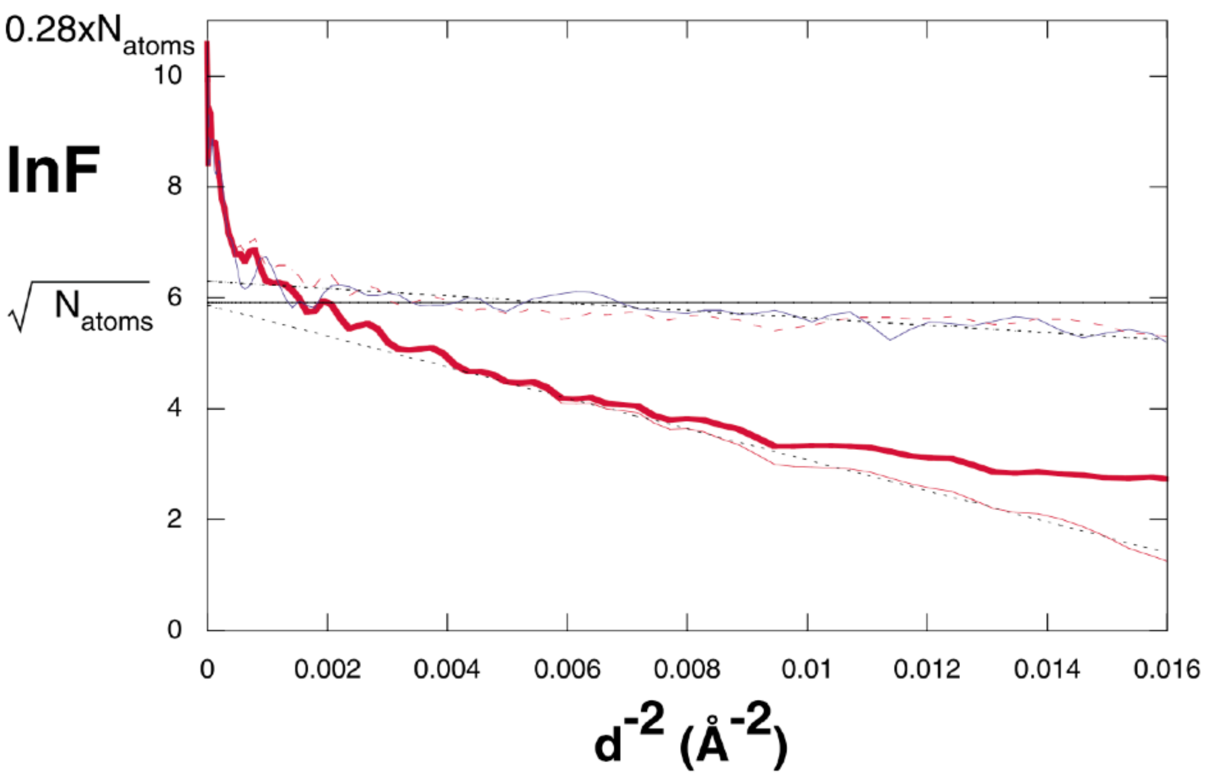


Tilt pairs are low resolution / structure at high

Global resolution and sharpening

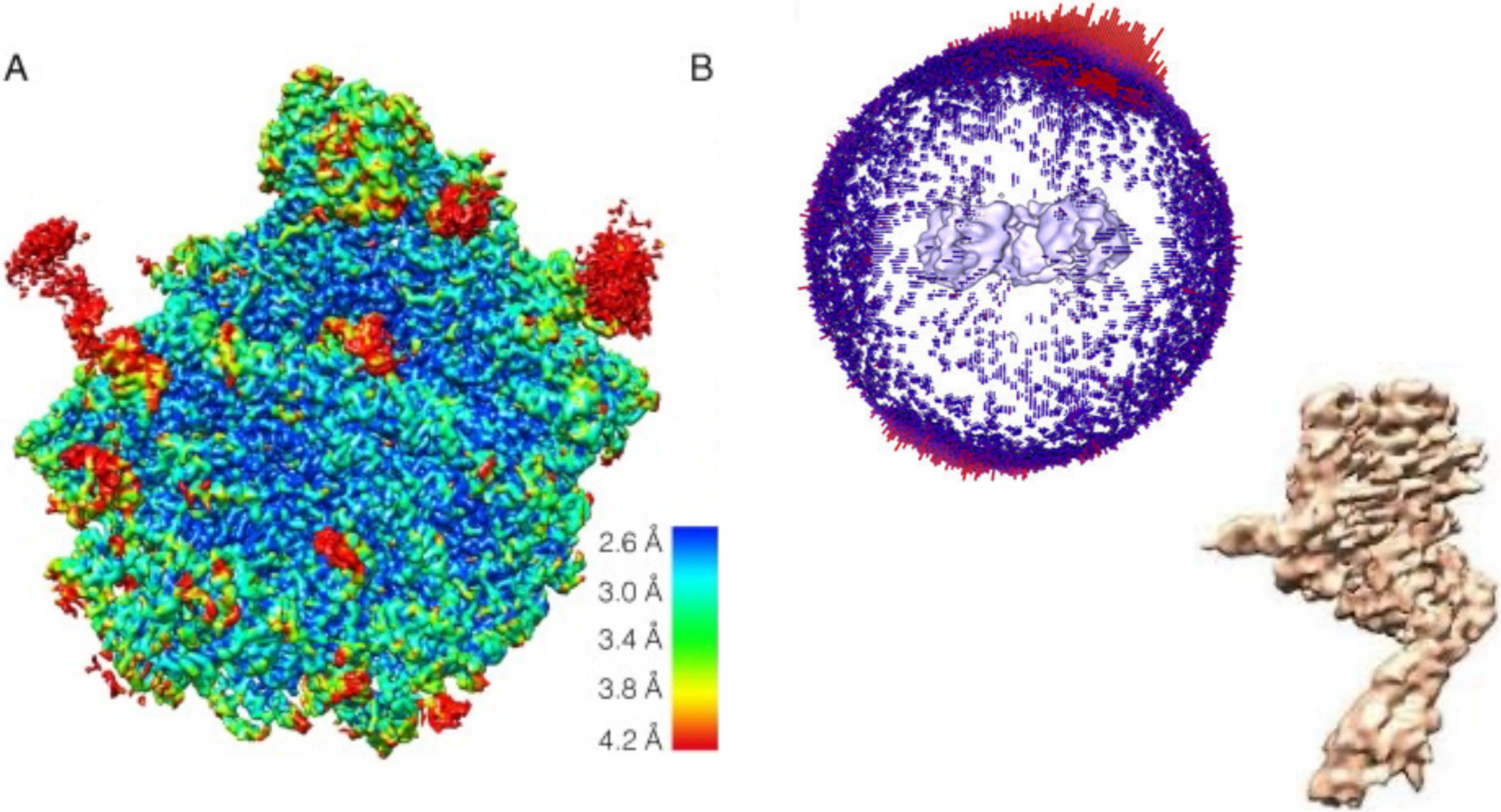
$$FSC(r) = \frac{\sum_{r_i \in r} F_1(r_i) \cdot F_2(r_i)^*}{\sqrt[2]{\sum_{r_i \in r} |F_1(r_i)|^2 \cdot \sum_{r_i \in r} |F_2(r_i)|^2}}$$





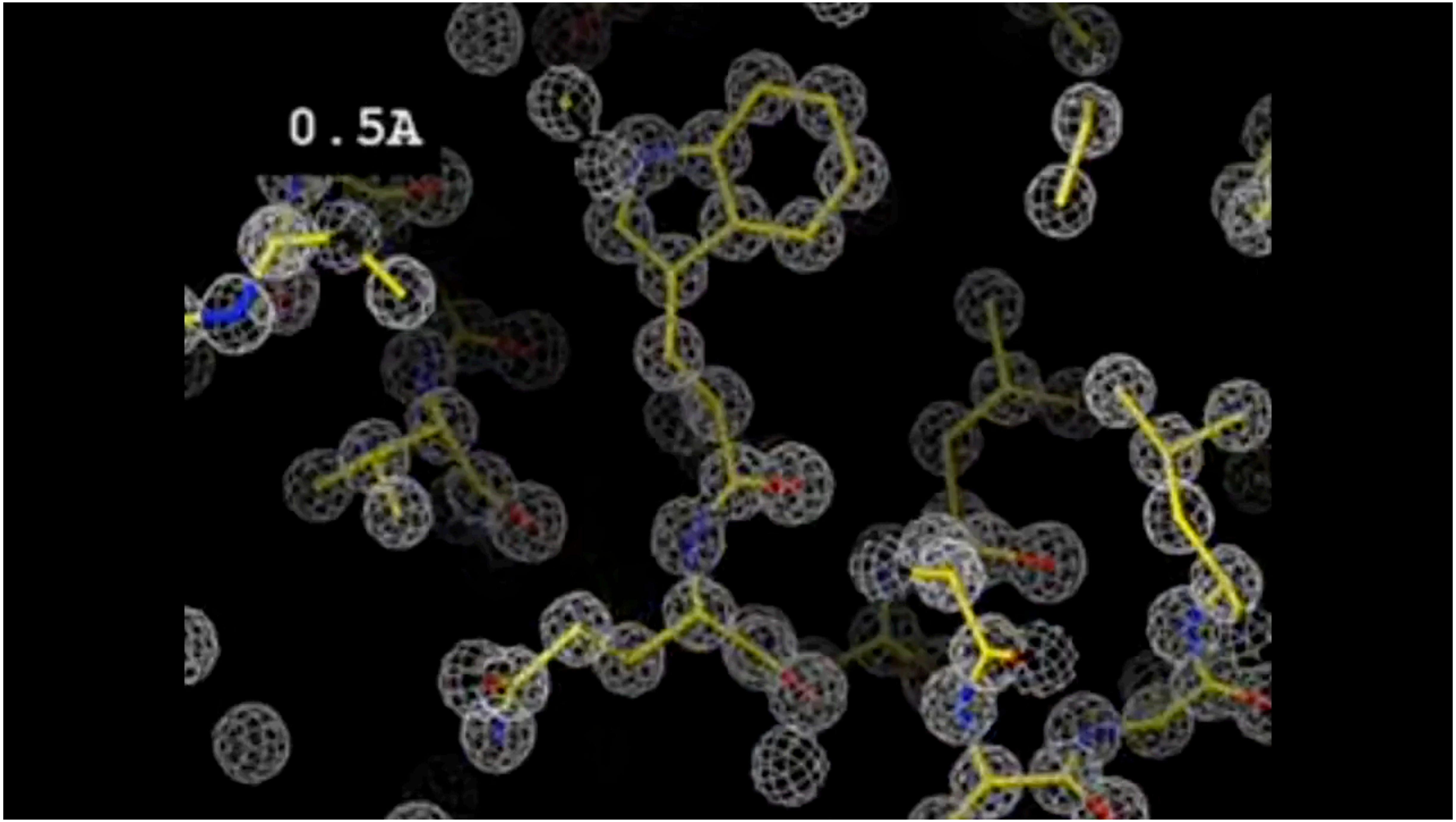
Limit interpretation to their consistent resolution

Local and directional resolution



Artefacts must be removed before interpretation

Biological chemistry



Biological chemistry is your best validation

Many thanks to Kyle Morris for the invitation to talk, Colin Palmer, Tom Burnley, and Helen Saibil for kindly providing diagrams for slides, and to everyone else for listening!

wellcome

Imperial College London

THE ROYAL SOCIETY